A SERIES SOLUTION FRAMEWORK FOR FINITE-TIME OPTIMAL FEEDBACK CONTROL, H-INFINITY CONTROL AND GAMES

A Dissertation

by

RAJNISH SHARMA

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 2008

Major Subject: Aerospace Engineering

A SERIES SOLUTION FRAMEWORK FOR FINITE-TIME OPTIMAL FEEDBACK CONTROL, H-INFINITY CONTROL AND GAMES

A Dissertation

by

RAJNISH SHARMA

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Co-Chairs of Committee,	Srinivas R. Vadali
	John E. Hurtado
Committee Members,	John L. Junkins
	D. Swaroop
Head of Department,	Helen Reed

December 2008

Major Subject: Aerospace Engineering

ABSTRACT

A Series Solution Framework for Finite-time Optimal Feedback Control, H-infinity Control and Games. (December 2008)

Rajnish Sharma, B.Tech.; M.Tech., Indian Institute of Technology, Kanpur

Co-Chairs of Advisory Committee: Dr. Srinivas R. Vadali
Dr. John E. Hurtado

The Bolza-form of the finite-time constrained optimal control problem leads to the Hamilton-Jacobi-Bellman (HJB) equation with terminal boundary conditions and to-be-determined parameters. In general, it is a formidable task to obtain analytical and/or numerical solutions to the HJB equation. This dissertation presents two novel polynomial expansion methodologies for solving optimal feedback control problems for a class of polynomial nonlinear dynamical systems with terminal constraints. The first approach uses the concept of higher-order series expansion methods. Specifically, the Series Solution Method (SSM) utilizes a polynomial series expansion of the cost-to-go function with time-dependent coefficient gains that operate on the state variables and constraint Lagrange multipliers. A significant accomplishment of the dissertation is that the new approach allows for a systematic procedure to generate optimal feedback control laws that exactly satisfy various types of nonlinear terminal constraints.

The second approach, based on modified Galerkin techniques for the solution of terminally constrained optimal control problems, is also developed in this dissertation.

Depending on the time-interval, nonlinearity of the system, and the terminal constraints, the accuracy and the domain of convergence of the algorithm can be related to the order of truncation of the functional form of the optimal cost function. In order to limit the order of the expansion and still retain improved midcourse performance, a waypoint scheme is developed. The waypoint scheme has the dual advantages of reducing computational efforts and gain-storage requirements. This is especially true for autonomous systems. To illustrate the theoretical developments, several aerospace application-oriented examples are presented, including a minimum-fuel orbit transfer problem.

Finally, the series solution method is applied to the solution of a class of partial differential equations that arise in robust control and differential games. Generally, these problems lead to the Hamilton-Jacobi-Isaacs (HJI) equation. A method is presented that allows this partial differential equation to be solved using the structured series solution approach. A detailed investigation, with several numerical examples, is presented on the Nash and Pareto-optimal nonlinear feedback solutions with a general terminal payoff. Other significant applications are also discussed for one-dimensional problems with control inequality constraints and parametric optimization.

DEDICATION

To my parents

ACKNOWLEDGEMENTS

I would like to thank my committee co-chairs, Dr. Srinivas R. Vadali and Dr. John. E. Hurtado, and my committee members, Dr. John Junkins and Dr. D. Swaroop, for their excellent guidance and support throughout the course of this research and my personal development. I am also thankful to Dr. Datta for acting as a substitute for Dr. Swaroop.

Thanks also go to my friends and colleagues and the faculty and staff of the Department of Aerospace Engineering for making my time spent at Texas A&M University a great experience.

Thanks and appreciations go to my mother, father and sisters for their encouragement, patience and love.

TABLE OF CONTENTS

CHAPTER Page

CHAPTER Page

LIST OF FIGURES

FIGURE Page

FIGURE Page

FIGURE Page

FIGURE Page

LIST OF TABLES

CHAPTER I

INTRODUCTION

Engineering processes can often be represented mathematically as dynamical systems and associated constraints. Control theory can be employed to obtain inputs to transfer the initial state of a dynamical system to the desired final condition. If the desired task is not only the satisfaction of the governing constraints but also the minimization/maximization of a given performance index, then a systematic approach to control design can be devised with the use of optimal control theory.

In general, the optimal control problem (OCP) is formulated to obtain the best strategy for minimizing or maximizing a performance index subject to the constraints imposed by the given dynamical system, actuation limits, and endpoints. The classical theory of calculus of variations[1] can be used to derive the necessary conditions for optimality[2] in the form of a two-point boundary value problem (TPBVP). The TPBVP can be solved by using standard direct or indirect methods, resulting in an open-loop candidate optimal control history. The open-loop control solution so obtained has to be subjected to the control history and the second order sufficient conditions before it can be declared to be locally optimal. In the "real-world" environment, with uncertainties in the initial conditions and other unmodeled effects, a locally optimal feedback solution is more useful than an open-loop solution, which needs to be recomputed for a new initial

---

This dissertation follows the style of *Journal of Guidance, Navigation and Control.*

condition. The methodology for obtaining optimal feedback controls is quite well developed for linear systems with linear desired end conditions. However, if the system or constraints, or both are nonlinear, then solving the OCP in a feedback setting can be extremely challenging to solve. A field of extremals can be generated either by solving the TPBVP for a number of initial and final conditions or, alternatively, by solving the Hamilton-Jacobi-Bellman (HJB) equation[3], whose solution satisfies both the necessary and sufficient conditions for optimality.

Unfortunately, even for relatively low dimensional nonlinear systems, solving the HJB equation with terminal constraints is a formidable task and closed-form solutions are available only for a few special cases. Furthermore, the feedback control design problem becomes more complicated if additional issues such as controller robustness and effective design for the worst-case disturbance and/or uncertain initial conditions are considered. Dynamic programming methods[4] in conjunction with the H-infinity formulation can be suitably utilized to obtain solutions for such problems.

Recently, there has been a renewed interest in solving pursuit-evasion problems, which can be cast in the framework of the theory of differential dynamic games[5]. Differential game problems can be formulated in many settings of interest to the aerospace community, namely, formation flying, air-traffic control, uncertainty modeling, circumnavigation and visual identification, collision avoidance, etc. The pioneering work of Isaacs[6] has led to the development of the Hamilton-Jacobi-Isaacs[7] (HJI) equation, which, like the HJB equation, is a first order nonlinear partial differential

equation. Analytical solutions to the HJI equation cannot be found for most cases of practical interest and one has to rely on numerical or approximate techniques.

This dissertation develops a general, Series Solution Method (SSM) for HJB/HJI equations, formulated for a class of nonlinear dynamical systems with nonlinear terminal constraints. The Galerkin Approximation Technique (GAT) with the use of orthogonal polynomial bases functions is discussed in detail. The solution to the HJI equation is examined in the context of non-cooperative[8], linear quadratic (LQ) and nonlinear cooperative games[9]. Preliminary analyses and LQ problems with control constraints are discussed. Finally, the solution to parametric optimization for dynamical systems and free final-time optimal control problem in the HJB settings is presented. The next sections of this chapter are focused on the relevant background and the organization of the dissertation.

1.1 Literature Survey

In this dissertation, the focus is on the solution of the HJB/HJI equations that arise in the field of optimal control of aerospace systems. Much of the previous work has been focused on the solution of the infinite-horizon problem.

For infinite-horizon OCP, A'lbrekht[11] used a series representation of the closed-loop cost function (also a Lyapunov function) to construct optimal feedback controls for nonlinear systems. Lukes[12] formalized the power series approach by expanding the cost and control functions about the origin. Dabbous and Ahmed[13] applied the series solution

method for optimal regulation of the angular momentum of a satellite. Carrington and Junkins[14] used the series expansion of the costate vector to solve the optimal finite-horizon spacecraft attitude regulation problem in feedback form. Yoshida and Loparo[15] presented a functional approximation theory for the optimal regulation of nonlinear systems. Garrard et al.[16] presented an approximation method by neglecting the time derivatives of certain nonlinear terms in the HJB equation. Cloutier[17] introduced the State Dependent Riccati Equation (SDRE) approach for the approximate solution of nonlinear OCP, a concept borrowed from the LQ framework. Tewari[18] and Sharma and Tewari[19] presented iterative approximation techniques, which use specific forms of a positive definite cost function, for the spacecraft attitude control and tracking problems. Navasca and Krener[20] developed a method by mixing the power series approach with Pontryagin's principle[21] to solve higher-order infinite time HJB equation. In the areas of trajectory optimization and motion planning, Cerven and Bullo[22] utilized higher-order series expansion method[23] which can be applicable to special classes of optimal control problems, e.g., infinite time horizon, linear terminal constraints, etc. Lyshevski[24] presented an approximate method for designing bounded controllers via generalized nonquadratic functionals.

Several other approximation methods have also been developed in the literature. The method of "Approximating Sequence of Riccati Equations" (ASRE) was introduced by Cimen and Banks[25]. This iterative method approximates the feedback solution to nonlinear optimal control problems by representing the nonlinear system and non-quadratic performance index, respectively, as a sequence of linear systems and quadratic

cost functionals about a sequence of nominal solutions. The method is not directly applicable, however, to problems with terminal constraints. Xin and Balakrishnan[26] developed the $\theta - D$ suboptimal approach, which is based on the addition of perturbations to the original cost function until an exact closed-form solution to a perturbed HJB equation can be achieved for a neighboring problem. Beard et al.[27-28] developed a Galerkin successive approximation technique that solves a sequence of linear Generalized HJB (GHJB) equations. In a compact domain, for analytic dynamical systems, the iterative process converges to the solution of the HJB equation if the process is started with a stabilizing control. To solve infinite-time robust optimal control problems, extensions of the GHJB methodology can be found in Beard and McLain[29]. Park and Tsiotras[30-31] used the wavelet basis functions to solve the GHJB equation using the iterative Galerkin procedure together with a collocation approach.

Richardson and Wang[32] have applied a finite-volume technique to solve the HJB equation and provide an example involving three states and three controls with control constraints; they did not address the problem of terminal constraints. Yet another method is the level-set method of Jin & Osher[33]. This method uses the gradient of the Hamilton-Jacobi equations which is solved by using finite difference schemes, essentially non-oscillatory (ENO) and weighted essentially non-oscillatory (WENO) methods of discretization. These methods can handle sharp discontinuities in the value function Based on finite difference schemes which are not guaranteed to converge to optimal solution, Huang and Lu[34] presented some alternatives to solve infinite time HJB/HJI equations. The methods of frozen Riccati Equations and nonlinear matrix inequalities

(NLMI) in the form of finite differences are applied to solve infinite time HJB equations. Related research exists on the theory of viscosity solutions[35-36], stabilizing solutions of the HJB equations for which conventional differentiable solutions do not exist. Much theory with examples is given in Ref. [36].

As discussed above, much of the work has been focused on the infinite-time regulation problem. Notably, the fixed final-time OCP with nonlinear terminal constraints has not received its due attention. The primary reason is the need to deal with hard terminal constraints and the associated computation of the constraint Lagrange multiplier. A special case of the hard terminal constraint problem is one where the final states are specified a priori or can be uniquely determined from the constraint equation. Recently, Guibout and Scheeres[37] and Park and Scheeres[38] developed a series solution method using canonical transformations and generating functions. Different types of generating functions can be chosen depending on the class of problems (boundary conditions) being solved. A generating function of one type can be converted into that of another type via the Legendre transformation. A special computational technique was developed to generate closed-form solutions for a class of TPBVPs arising in the context of impulsive orbit transfer[39], formation flying, and other OCPs. The works referred above do not treat OCPs for systems with explicit nonlinear terminal constraints.

Applications of the series solution method to the Hamilton-Jacobi-Isaacs (HJI) equation can be found in Huang and Lin[40-41] and Tsiotras et al.[42] Beard and McLain[29] and Wise and Sedwick[43] utilized successive approximation approaches to solve the HJI equation. Khalaf and Lewis[44-46] presented a neural network-based method[47] of solution

for the HJI for constrained input nonlinear systems. Soravia[48] addressed H-infinity control of nonlinear systems by using viscosity solutions. van der Schaft[49] presented a treatment of the geometric nature of the stabilizing solution of the HJI equation by exploiting symplectic geometry of integrable systems using Hamiltonian perturbation theory[50]. However, all of these methods require further extensions to handle problems with nonlinear terminal constraints and/or a finite time horizon.

Worst case design scenarios[51] provide interesting applications of the HJI equation. It is shown in Ref. [45] that H_2 / H_∞ control problems can also be posed as zero-sum[52] two-player differential games. The standard H_2 / H_∞ formulation in state-space form, for linear systems, is given by Doyle et.al[53], which requires the existence of solutions of a set of Riccati equations[54]. Extensions of the theory to nonlinear systems can be found in Refs. [49, 55].

Although many elegant aspects of the two-player, zero-sum game have been addressed over the years, there are still some research questions that must be answered. Various open-loop and feedback methods for pursuit-evasion[56] problems such as feedback linearization[57], semi-direct methods[58] and well-known indirect optimization methods[59] exist. However, none of these methods provide a straightforward approach to obtain a higher-order feedback solution to nonlinear dynamic game problems due to the inherent difficulties associated with the hyperbolic nature of the cost function, discontinuous minimizing solution and the conditions associated with solving the HJI equation and the investigation of the saddle point region.

For the two-player, zero-sum games with Nash solutions[60-62], in the LQ setting, much of the theory involving a single pursuer and a single evader was established by Isaacs[6-7]. Significantly less research, however, has been accomplished for the cooperative multiplayer situation. Many facets of the cooperative linear quadratic Pareto optimality[9] game have been widely studied. Pareto optimality is concerned with the best joint decisions among players. This theory has not been extended to obtain Pareto optimal feedback solutions of nonlinear cooperative games[61]. Solution of the finite-horizon HJI equation with hard terminal constraints and its applications to cooperative/non-cooperative dynamic games[63] and H-infinity design problems have not been obtained, so far.

1.2 Motivation and Dissertation Overview

A primary objective of this research work is to develop a structured approach for solving the finite-time HJB equation for a class of polynomial nonlinear systems with nonlinear terminal constraints. Steps similar to those developed for the HJB equation can be followed to solve a class of robust control problems governed by the H-infinity formulation and nonlinear differential games. The development of a procedure to solve the finite-time HJI equation with a general terminal constraint is another important extension of the work. The class of problem mentioned above can be solved using SSM developed in this dissertation.

The outline of the dissertation is as follows:

The next chapter begins with the essential background material needed for formulating an optimal control problem. A brief overview of the first order necessary conditions for optimality in the context of an open-loop solution and the development of the HJB equation are presented. Important results which connect classical calculus of variations approach to dynamic programming approach are treated next. An LQ example is presented to address the motivating factors behind SSM. The next section is devoted to the presentation of the key contribution of this dissertation—SSM, developed to solve optimal control problems for smooth, analytic, polynomial dynamical systems with affine control. Subsequent sections of this chapter discuss the applicability of SSM to OCPs involving polynomial systems and the investigation of another novel method for solving the finite-time HJB equation by the use of Galerkin techniques. The SSM for terminal constraints is used throughout the subsequent chapters.

Chapter III specifically focuses on the application of SSM for obtaining feedback control solutions of several higher-order optimal control problems, including spacecraft detumbling and low-thrust orbit transfer examples. The solution obtained for each problem is benchmarked against its corresponding open-loop solution to ascertain the accuracy of SSM within the order of truncation. Sensitivities of the feedback solutions as well as the cost-to-go are evaluated with respect to perturbations in the known initial states. It is shown that for small perturbations, highly accurate guidance solutions can be achieved by using a third-order feedback control law.

Chapter IV discusses an innovative idea of blending the notion of a waypoint scheme with SSM for solving extended-horizon HJB equations in the context of

designing optimal feedback control laws for nonlinear systems with terminal constraints. The technique addresses a need for enlarging the domain of convergence of the series solution method. The concept of waypoints is introduced by partitioning the overall time interval of the given problem into smaller segments and the series solution method is applied within each segment with the use of stored gains, computed for one segment only. The updated methodology is applied to highly nonlinear systems including a minimum-fuel Earth to Mars orbit transfer problem. Several examples are demonstrated and the results are compared with the corresponding open-loop solutions to demonstrate the efficacy of the proposed method.

In chapter V, a novel application of SSM for solving terminally constrained, the finite-time Hamilton-Jacobi-Isaacs (HJI) equation is presented. The first example considered is the problem of nonlinear differential games in orbits. The resulting HJI equation is solved in order to construct nonlinear feedback strategies for finite-time pursuit and evasion scenarios involving space assets.

Chapter VI is devoted to the treatment of pursuit-evasion examples of two-player nonlinear differential games involving orbiting satellites, cooperative games and nonlinear Pareto-optimal feedback solutions.

Chapter VII is devoted to the extended applications of series solution methods to deal with control constraints and parameter optimization problems.

Finally, chapter VIII concludes the summary of the work along with the remarks on each chapter.

CHAPTER II

POLYNOMIAL SERIES EXPANSION METHODOLOGIES

In this work, the Polynomial Series Expansion Methodology is taken to mean a body of methods in which the cost-to-go function, also known as the optimal return function, is written as a series of terms, each generally containing products of unknown time-dependent control gains, the system states, and terminal Lagrange multipliers. Two particular polynomial expansion methods that will be discussed in this chapter are a Series Solution Method (SSM) and a Galerkin Approximation Technique (GAT). There are some remarkable differences between these two polynomial expansion methods.

1. As the name suggests, SSM approximates the cost-to-go function using a power series expansion in the system states and terminal Lagrange multipliers, whereas the GAT uses an expansion in the form of orthogonal polynomial bases functions of the system states and terminal Lagrange multipliers.

2. The time-dependent control gains in each method are governed by ordinary differential equations. The ordinary differential equations are found by matching coefficients on both sides of the HJB equation in SSM; GAT finds the ordinary differential equations by minimizing the residual error in satisfying the HJB equation. Moreover, the ordinary differential equations obtained from SSM are

uncoupled, linear equations after the first level, whereas all the ordinary differential equations produced by GAT are coupled and nonlinear.

3. Finally, though the domain of convergence of the feedback methods is dependent on the number and type of the basis functions utilized in the expansion, SSM achieves high local accuracy and it is ideally suited for systems inherently consisting of polynomial nonlinearities; whereas GAT is able to accommodate a larger class of nonlinearities and has a wider domain of convergence with a uniform global error in the HJB solution. These issues will be discussed in more detail in the following sections.

This chapter begins with an introduction to the classic optimal control problem in section 2.1. Section 2.2 presents the necessary conditions that must be satisfied by the solution to the optimal control problem. Included in this portion is a new, important result of the terminally constrained problems. Section 2.3 revisits the classic linear quadratic optimal control problem subject to terminal constraints. This classic problem is viewed from a new perspective, and although the end result is the same as that gained from the traditional view, the benefit is an obvious natural extension for nonlinear problems. This natural extension, SSM, is presented in section 2.4; examples illustrating the application of SSM are presented in sections 2.5 and 2.6. Section 2.7 presents the theory underlying GAT, and includes a few illustrative examples. Finally, section 2.8 compares and contrasts the performances of SSM and GAT.

2.1 Optimal Control Problem

The traditional optimal control problem can be written using the Bolza form as

Minimize:

$$J = \phi(x(t_f)) + \int_{t_0}^{t_f} L(x,u,\tau)d\tau \tag{2.1}$$

for a nonlinear dynamical system:

$$\begin{aligned} &\dot{x} = f(x,u,t); x(t_0) = x_0; \\ &\Psi \equiv \psi(x(t_f)) - \psi_f = 0;\ \psi, \psi_f \in \Re^{p \le n} \end{aligned} \tag{2.2}$$

where $f : \Re^{n+m+1} \to \Re^n$ is a smooth, analytic, vector-valued function with the state, $x \in \Re^n$ and the control, $u \in \Re^m$. The initial time t_0, the terminal time t_f, and the initial condition on the state vector x_0, are assumed to be given. It is also assumed that $f(0,0,t) = 0$ and the affine control vector u is unconstrained. At the terminal time, the constraints can either be soft or/and hard. A soft constraint is a scalar function of $x(t_f)$ and is given by $\phi(x(t_f))$. A terminal constraint can be vector valued and is given by Ψ in Eq. (2.2). The value of the constraint is defined as ψ_f, to satisfy the condition $\psi(x(t_f)) = \psi_f$. In the optimal control literature, such a terminal constraint is also known as the hard constraint which could be specified as an explicit point constraint or in a more general way; as a nonlinear hyper-surface. The cost-to-go function is defined by replacing the lower limit of integration with the current time in Eq. (2.1) .It is termed as the value function in the

dynamic programming literature. Further assumptions are that the value function and the controls are smooth with respect to the states, Lagrange multipliers, and time.

2.2 Derivation of Important Results

In this section, the important necessary conditions that must be satisfied by the solution to the optimal control problem are presented. These necessary conditions are derived using two different approaches. The first is a variational approach and the second is a dynamic programming approach. Each approach, of course, ultimately leads to the same optimal solution. The variational approach requires the solution of a vector of costate ordinary differential equations, whereas the dynamic programming[4] approach requires the solution of a single partial differential equation (PDE) for the cost-to-go. Eventually this PDE must also be solved via discretization, except for simple problems for which analytical solutions exist.

The terminal constraint and Eq. (2.2) are augmented to the cost function and the augmented cost function is written as

$$J_a = \phi(x(t_f)) + \nu^T[\psi(x(t_f)) - \psi_f] + \int_{t_0}^{t_f} \{L(x,u,t) + \lambda^T(f(x,u,t) - \dot{x})\}dt \qquad (2.3)$$

where, ν is a vector of constant Lagrange multipliers that enforce the terminal constraints and λ is the costate vector. The necessary conditions for optimality[64] can be derived by defining the Hamiltonian, H:

$$H = L(x,u,t) + \lambda^T f(x,u,t) \qquad (2.4)$$

The first variation of the augmented cost function is written as

$$\delta J_a = \left[\frac{\partial \phi}{\partial x(t_f)} + \frac{\partial \psi}{\partial x(t_f)}^T \nu - \lambda(t_f) \right]^T \delta x(t_f) + \left[\psi(x(t_f)) - \psi_f \right]^T d\nu$$
$$+ \int_{t_0}^{t_f} \left[\left(\frac{\partial H}{\partial x} + \dot{\lambda} \right)^T \delta x + \left(\frac{\partial H}{\partial \lambda} - \dot{x} \right)^T \delta \lambda + \left(\frac{\partial H}{\partial u} \right)^T \delta u \right] dt \tag{2.5}$$

First-order necessary conditions that must be satisfied by the optimal state and control are obtained by requiring that the first variation δJ_a in Eq. (2.5) vanish for admissible variations in x, u, λ, and ν. The necessary conditions for optimizing the augmented performance index are[64]:

$$\dot{x} = \frac{\partial H}{\partial \lambda}$$
$$\dot{\lambda} = -\frac{\partial H}{\partial x} \tag{2.6}$$

$$u^*(t) = \min_u (H) \tag{2.7}$$

where $u^*(t)$ is the optimal control.

The transversality condition[65] is

$$\lambda(t_f) = \frac{\partial \phi(x(t_f))}{\partial x(t_f)} + \left(\frac{\partial \psi(x(t_f))}{\partial x(t_f)} \right)^T \nu \tag{2.8}$$

A further obvious necessary condition is that the terminal constraint must be satisfied.

A candidate open-loop solution to the above OCP must satisfy the first-order necessary conditions derived as Eqs. (2.6)-(2.8). The candidate solution must also satisfy the second-order necessary conditions, namely, the convexity, normality, and the Jacobi conditions[3] before being accepted as the required solution. Note that the open-loop

solution, if it can be obtained, satisfies the given initial condition only; it has to be recomputed for a different initial or terminal condition. Typically, an open-loop solution is utilized as a benchmark solution to verify other approximate solution techniques. Also, it can be used as a nominal to perturbation control or neighboring optimal feedback control problems.

To determine the optimal control in feedback form, a different approach via dynamic programming formalism can be considered. As for the calculus of variations approach, it also provides necessary and sufficient conditions for the above minimization problem in the form of the HJB equation and the associated boundary conditions. Explicit feedback control laws can be obtained from the solution to the HJB equation, a nonlinear partial differential equation for the optimal value function (cost-to-go) J^* as shown below:

$$\frac{\partial J^*}{\partial t} = -H(x(t), u^*(t), \frac{\partial J^*}{\partial x}, t) \tag{2.9}$$

with the boundary condition

$$J^*(x(t_f), t_f) = \phi(x(t_f)); x(t_f) \text{ such that } \psi(x(t_f)) = \psi_f \tag{2.10}$$

The key result that connects dynamic programming and the calculus of variation approach is the relationship between the costate vector and the gradient of the value function:

$$\lambda(t) = \frac{\partial J^*(x(t), t)}{\partial x(t)} \tag{2.11}$$

Indeed, using the Lagrange multiplier rule, Eq. (2.10) can also be written as

$$J^*(x(t_f),t_f) = \phi(x(t_f)) + \nu^T[\psi(x(t_f)) - \psi_f] \tag{2.12}$$

and the partial derivative of the above expression w.r.t. $x(t_f)$ results in the boundary condition of Eq. (2.8).

$$\frac{\partial J^*(x(t_f),t_f)}{\partial \nu} = (\psi(x(t_f)) - \psi_f) = 0 \tag{2.13}$$

A detailed treatment of the boundary conditions can be found in Dreyfus[1-2].

Another key sensitivity result can be arrived at by considering the effect of the differential in the constraint Lagrange multiplier vector on the cost, in the neighborhood of the optimal solution. It can be observed from Eq. (2.5) that if all of the necessary conditions are satisfied, then the gradient with respect to ν, of the augmented cost function, along the optimal trajectory is

$$\frac{\partial J_a}{\partial \nu} = \frac{\partial J(x(t_0),t_0)}{\partial \nu} + (\psi(x(t_f)) - \psi_f) = 0 \tag{2.14}$$

Since the terminal conditions are also satisfied as part of the necessary conditions, and since the initial time is arbitrary, the above result can be generalized and also stated as

$$\frac{\partial J^*(x(t),t)}{\partial \nu} = 0 \tag{2.15}$$

Equation (2.15) is the main result which will be exploited throughout the work. This important result has not been utilized in any of the standard references found in the literature for obtaining numerical solution to terminally-constrained OCPs. It is also important to point out that unlike the solution obtained from a two-point boundary-value problem, a solution that satisfies the HJB equation, if it exists, satisfies both the necessary and sufficient conditions for a minimum.

The partial derivative of Equation (2.12) with respect to ν shows that Equation (2.15) is also satisfied at the final time:

$$\frac{\partial J^*(x(t_f),t_f)}{\partial \nu} = [\psi(x(t_f)) - \psi_f] = 0 \tag{2.16}$$

The above equation provides a stationary condition to obtain an extremum of the cost-to-go function with respect to the terminal Lagrange multiplier. For LQ problems, it can be analytically verified that J^* attains the maximum value at the solution of Eq. (2.16). For nonlinear problems, this duality can be numerically checked within the validity of the solution. The second order conditions are based on the existence of the solutions of the accessory minimum problem (AMP) which is an approximation of the HJB equation.

2.3 An Alternative Derivation of the LQ Terminal Controllers

The Linear-Quadratic (LQ) problem is a special case for which there is an exact, closed-form feedback solution for the HJB equation, with the control gains resulting from the solution to a Riccati equation and a set of linear auxiliary differential equations. In this section, the classic linear quadratic optimal control problem subject to terminal constraints is revisited by using the dynamic programming approach[3]. When viewed from a new perspective, we see that an extension to handle nonlinear problems is revealed.

Consider the following LQ problem with known initial and final times:

Minimize:

$$J = \frac{1}{2}\int_{t_0}^{t_f} \left[x^T Q x + u^T R u \right] dt \tag{2.17}$$

subject to

$$\dot{x} = Ax + Bu \tag{2.18}$$

and a linear terminal constraint

$$Cx(t_f) = \psi_f \tag{2.19}$$

Before solving the optimal control problem, the following assumptions are made,

i. Q is symmetric and positive semi-definite.

ii. The control-weighting matrix R is symmetric and positive-definite.

iii. Controllability of the pair (A, B) and observability of the pair $(A, \sqrt{Q})$ exist.

It is noted that when the number of constraints is equal to the number of states and the matrix C in Eq. (2.19) is invertible, then the final state can be explicitly solved for. For cases where the number of constraints is less than the dimension of the state vector, the final state is determined as a result of the optimization process.

The Hamiltonian for the problem is

$$H = \frac{1}{2}\left[x^T Q x + u^T R u \right] + \lambda^T \left[Ax + Bu \right] \tag{2.20}$$

The application of the optimality condition results in the following relationship:

$$u = -R^{-1} B^T \lambda \tag{2.21}$$

Substitution of Eq. (2.21) in Eq. (2.20) results in the following form of the Hamiltonian:

$$H = \frac{1}{2}\left[x^T Q x - \lambda^T B R^{-1} B^T \lambda \right] + \lambda^T \left[Ax \right] \tag{2.22}$$

The transversality condition is written as

$$\lambda(t_f) = \frac{\partial J^*(x(t_f), t_f)}{\partial x(t_f)} = C^T \nu \tag{2.23}$$

The next step for obtaining a feedback control law is an expansion of the cost-to-go. Ideally, it is desirable to express the cost function in terms of the current and final states. However, this is not possible directly, since an expansion of this type does not lead to explicit terminal boundary conditions for the gains. This situation can be identified with the lack of identity transformations for certain types of generating functions in the canonical transformation approach. Since the Hamiltonian is quadratic in x and λ and Eq. (2.23) suggests a linear dependence of λ on ν, the following quadratic form is written for $J^*(x(t), t)$:

$$J^*(x(t), t) = \frac{1}{2} x(t)^T S(t) x(t) + \nu^T K(t) x(t) + \frac{1}{2} \nu^T P(t) \nu - \nu^T \psi_f \tag{2.24}$$

Equation (2.24) is a modified version of a similar form given in Bryson and Ho[26]. The addition of the last term in Eq. (2.24) is motivated by its presence in Eq. (2.3). The matrices $S, K,$ and P in the above equations are time-varying gain matrices of appropriate dimensions. Differential equations and the appropriate boundary conditions for the gains have to be determined. Note that *S* and *P* can be assumed to be symmetric.

The following results are obtained by utilizing Eqs. (2.11) and (2.15) for the LQ problem:

$$\lambda = \frac{\partial J}{\partial x} = Sx + K^T \nu \tag{2.25}$$

$$\frac{\partial J}{\partial \nu} = 0 = Kx + P\nu - \psi_f \tag{2.26}$$

The boundary conditions for the above gains can be obtained by inspecting Eqs. (2.19) and(2.23), as given below:

$$S(t_f) = 0 \tag{2.27}$$

$$K(t_f) = C \tag{2.28}$$

$$P(t_f) = 0 \tag{2.29}$$

Note also that these boundary conditions result in the following intuitive result for the cost at the final time:

$$J(x(t_f), t_f) = 0 \tag{2.30}$$

The HJB equation can be written as

$$\frac{\partial J}{\partial t} = -\frac{1}{2}\left[x^T Q x - \frac{\partial J}{\partial x}^T BR^{-1}B^T \frac{\partial J}{\partial x} \right] - \frac{1}{2}\frac{\partial J}{\partial x}^T [Ax] - \frac{1}{2} x^T A^T \frac{\partial J}{\partial x} \tag{2.31}$$

The differential equations for the gains are obtained by utilizing Eq. (2.24) in Eq. (2.31) and collecting the coefficients of the linear and quadratic terms involving x and ν. These equations are

$$\dot{S} = -Q - SA - A^T S + SBR^{-1}B^T S \tag{2.32}$$

$$\dot{K} = -K(A - BR^{-1}B^T S) \tag{2.33}$$

$$\dot{P} = KBR^{-1}B^T K^T \tag{2.34}$$

The optimal control law can be obtained from Eq.(2.21), via Eq. (2.25). However, in order to implement the control, the constant, ν must be calculated from Eq. (2.26) at any time, other than the final time, as follows:

$$\nu = -P(t)^{-1}(K(t)x(t) - \psi_f) \tag{2.35}$$

Sufficient conditions for a minimum, based on the existence of solutions to Eq. (2.32)-(2.34), can be found in Bryson and Ho[64]. Significance of the normality condition, $P(t) < 0$ for $t_0 \le t \le t_f$, can be seen from the evaluation of the value function by replacing ψ_f in Eq. (2.24) by using Eq. (2.26). The result is

$$J^*(x(t),t) = \frac{1}{2}x(t)^T S(t)x(t) - \frac{1}{2}\nu^T P(t)\nu \tag{2.36}$$

Since S is required to be the positive definite solution to the Riccati equation, the normality condition ensures that the value function remains positive definite, as it should be.

Another important observation for terminal Lagrange multiplier is that it also satisfies the dual properties as it maximizes the cost-to-go. It can be verified from the second derivative of Eq. (2.24) with respect to ν, which is,

$$\frac{\partial^2 J^*(x(t),t)}{\partial \nu^2} = P(t) \le 0 \forall t \tag{2.37}$$

In the absence of numerical errors, ν is a constant, irrespective of the time *t*, at which it is computed by using Eq. (2.35). The optimal control problem can be solved by forward integration of the state and costate differential equations, once the initial costate

vector is determined from Eq. (2.25). However, a feedback control implementation requires that the gains be stored and the control be implemented using the current state and time. In the feedback control implementation, there exists a singularity at the final time where ν cannot be calculated from Eq. (2.35). However, in practice, updating of ν is stopped just before the final time in order to avoid this problem.

The classical sweep method[64] designed for LQ problems with terminal constraints is a two step method, which considers two separate linear combinations of x and ν associated with time dependent gains for the costate and the value of constraint. On the other hand, the derivation of feedback control by using the dynamic programming approach is obtained by utilizing a single quadratic expansion only. As done for LQ problems, the backward or forward sweep method for nonlinear systems can be constructed as well. The process is more complicated and requires the use of symbolic manipulation and efficient exploitation of symmetry of the resulting tensors. The data storage burden can be reduced considerably if the symmetry pattern is identified. The symmetry pattern is readily identifiable in the cost-to-go expansion process (dynamic programming approach). A minimal polynomial expansion can be utilized with all possible combinations of like terms in the expansion considered only once.

In the next section, the procedure described above for the LQ problem is extended to the more general case of a nonlinear system with a nonlinear terminal constraint.

2.4 Series Solution Methodology for the Nonlinear Problem

Here, SSM for nonlinear systems is presented. The explicit form of the optimal value function, which is the solution of the HJB equation obtained via Taylor series approximation, provides the analytic optimal control via the relationship of Eq. (2.11). Series solution methods are attractive for polynomial systems which are analytic and have a dominating linear term. Herein, it is assumed that the states are nondimensionalized such that the convergence properties of power series expansion are satisfied. Alternatively, the nonlinear perturbation dynamics close to the nominal optimal trajectory can be considered, to justify the use of SSM. Required assumptions regarding smoothness of the value function and existence of solutions are also made.

Since the Hamiltonian is a function involving the states and costates, the cost-to-go function is expanded in a polynomial series involving x and ν as shown below:

$$\begin{aligned} &J^{*}(x(t),t) = S_{1_{ij}}(t)x_i x_j + S_{2_{ijk}}(t)x_i x_j x_k + S_{3_{ijkl}}(t)x_i x_j x_k x_l ... + \\ &K_{1_{pj}}(t)\nu_p x_j + K_{2_{pqj}}(t)\nu_p \nu_q x_j + K_{3_{pij}}(t)\nu_p x_i x_j + ... \\ &-(\psi_f)_p \nu_p + C_{1_{pq}}(t)\nu_p \nu_q + C_{2_{pqr}}(t)\nu_p \nu_q \nu_r + ... + h.o.t \\ &i, j, k, l .....\text{etc.} = 1, 2, 3, ..., n;\ p, q, r .....\text{etc.} = 1, 2, 3, ..., p \le n \end{aligned} \tag{2.38}$$

where $S_{\mathrm{i}}(t), R_{\mathrm{i}}(t), K_{\mathrm{i}}(t), C_{\mathrm{i}}(t)...$, i=1,2,3,.. , are gain tensors (expressed using indicial notation), having time-dependent elements. This polynomial expansion of the cost-to-go is at the heart of the method presented in this chapter.

The SSM requires a polynomial form of the dynamical system. If the system is not given in a polynomial form, it can also be expanded about a suitable reference

solution maintaining the properties of series convergence. For synthesizing a higher-order feedback law, the open-loop solution can be considered as a nominal solution. Differential equations for the gain elements can be obtained by substituting Eq. (2.38) into the HJB equation (Eq. (2.9)) and collecting the coefficients of like powered terms involving x and ν. For example, it can be seen that $S_1(t)$ satisfies the familiar Riccati equation. Terminal boundary conditions on the gain variables can be postulated by using Eqs. (2.8) and (2.16). Once the gain elements are obtained via backward integration, the costate vector can be determined from Eq. (2.11) and ν can be determined from Eq. (2.15) by using vector reversion of series[66].

Generation of the equations for the gain elements is a tedious process but it can be simplified by the use of symbolic manipulation programs like Mathematica or Maple. Structured approaches to SSM have been presented in Refs. [15] and [67]. It is convenient to combine x and ν into a single extended vector for generating the gain equations. The equations required in this work were generated by using the software package, Maple. However, the general development of the computer code based on indicial notation is described in Appendix A. The complete procedure for a symbolic tool for higher-dimensional systems is also discussed in Appendix A. In the next section, a closed-form solution to the HJB of a one dimensional system with a nonlinear constraint is presented and this solution is used to estimate the errors in the control and in the satisfaction of the HJB equation.

In order to check the validity of the obtained series solution of the optimal return function, $J^*(x(t),t)$, it must satisfy the properties of a Lyapunov function. The valid range of initial conditions can be found by checking the condition given as:

$$J^*(x(t_0),t_0) \geq 0 \tag{2.39}$$

2.5 Examples with Analytical Solutions

Two examples are presented in this section involving an LQ problem.

The first example treats the LQ problem with a quadratic terminal constraint and the second example considers the case of a cubic terminal constraint. An analytical solution not being readily available, this example is solved using SSM. For both the cases, it is assumed that the constraint is regular, i.e., at least one real solution exists for $x(t_f)$. These problems are quite simple if the constraint is solved for $x(t_f)$, a priori. However, solving for $x(t_f)$ is not practical for multiple constraints, posed in a higher dimensional space. It will be desirable for the algorithm to choose the appropriate final condition that satisfies the constraint and results in the least cost. As will be shown, the solution to the HJB equation does have the required properties and SSM also duplicates these results, albeit, approximately.

The example OCP is

$$\text{Minimize } J = \frac{1}{2}\int_0^{t_f} u^2 dt$$
$$\text{subject to } \dot{x} = u, \quad \psi = x(t_f) + \varepsilon x(t_f)^2 = \psi_f \ ; x(0), \psi_f, \text{and } t_f = \textit{given}. \tag{2.40}$$

where ε is a small parameter.

The detailed analytical solution of this simple problem, which demonstrates that the value function may not be smooth even for the simplest of problems, is presented in Ref. [73] by using SSM.

The HJB equation for the example problem is

$$\frac{\partial J^*(x(t),t)}{\partial t} = J_t^* = \frac{1}{2} J_x^{*2} \tag{2.41}$$

subject to

$$J_x^*(x(t_f), t_f) = \lambda(t_f) = \nu \psi_x(x(t_f)) = \nu(1 + 2\varepsilon x(t_f)) . \tag{2.42}$$

The analytical solution to the optimal value function is

$$J^*(x(t),t) = \nu \left[\frac{\psi + \frac{\nu}{2}(t - t_f)}{1 - 2\varepsilon(t - t_f)\nu} - \psi_f \right] \tag{2.43}$$

where the solution to ν can be obtained from the equations given below:

$$\begin{aligned} &[1 - 2\varepsilon(t - t_f)\nu]^2 = \frac{(1 + 2\varepsilon x)^2}{[1 + 4\varepsilon\psi_f]} \ ; \ \varepsilon \neq 0 \\ &\nu = -\left(\frac{x - \psi_f}{t - t_f} \right) \ ; \ \varepsilon = 0 \end{aligned} \tag{2.44}$$

The series solution to the above problem is given below:

$$\lambda = -u = \left[\nu + 2\varepsilon(t - t_f)\nu^2 + 4\varepsilon^2 (t - t_f)^2 \nu^3 + 8\varepsilon^3 (t - t_f)^3 \nu^4 ... \right] \psi_x(x(t)) \tag{2.45}$$

The above series is a binomial expansion which converges if $\left|2\varepsilon(t-t_f)\nu\right|<1$ and it is recognized that as more terms in the series are taken, the resulting control law takes the form:

$$u=-\nu\left[\frac{1}{1-2\varepsilon(t-t_f)\nu}\right]\psi_x=-\nu\left[\frac{1}{1-2\varepsilon(t-t_f)\nu}\right](1+2\varepsilon x) \tag{2.46}$$

The process of using series reversion yields a unique solution to ν:

$$\begin{aligned}\nu = F &- 3\varepsilon(t-t_f)F^2+10\varepsilon^2(t-t_f)^2F^3-35\varepsilon^3(t-t_f)^3F^4\\ &+126\varepsilon^4(t-t_f)^4F^5-462\varepsilon^5(t-t_f)^5F^6+\ldots..\ ;\end{aligned} \tag{2.47}$$

$$\text{where } F=-\frac{1}{(t-t_f)}\left(\frac{\psi(x(t))-\psi_f}{\psi_x^2(x(t))}\right)$$

Numerical values selected for this example are $\varepsilon=0.5$, t_f=5 sec, and $\psi_f=0.5$.

Figure (2.1) shows the exact, bi-modal cost-to-go function for this problem, evaluated for a range of initial conditions. The series solution matches the exact solution in the ranges indicated by '*' for six-term expansions of Eq. (2.45) and(2.47). For $x(0)$ in the range (-2, 0), the six-term series solution does not converge, i.e., the closed-loop system is not stable.

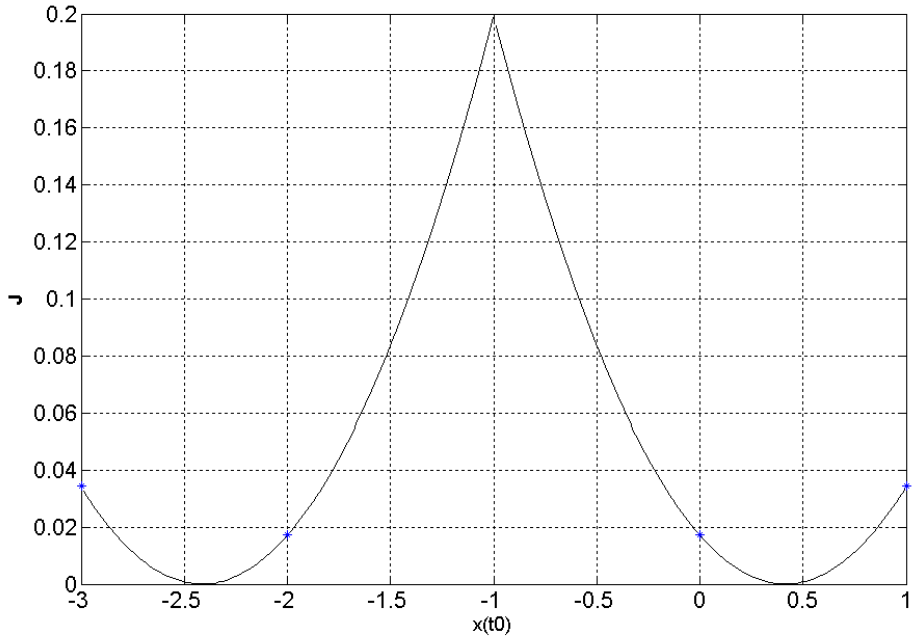


Fig. 2.1 Optimal Value Function: Comparison of the Results of the Series and Exact Solutions

To check the duality condition of terminal Lagrange multiplier given by Eq.(2.37), the second partial of Eq. (2.43) with respect to ν is obtained as,

$$\frac{\partial^2 J^*(x(t),t)}{\partial \nu^2} = J^*_{\nu\nu} = \frac{(t-t_f)(1+4\varepsilon\psi)}{[1-2\varepsilon\nu(t-t_f)]^3} \equiv \frac{(t-t_f)[1+2\varepsilon x(t)]^2}{[1-2\varepsilon\nu(t-t_f)]^3} \tag{2.48}$$

It is known that the series in Eq. (2.45) converges if $\left|2\varepsilon(t-t_f)\nu\right| < 1$; hence the denominator of Eq. (2.48) is always a positive quantity for all feasible solutions of ν in the above range. Hence, the duality relation,

$$J^*_{\nu\nu}(x(t),t) \le 0 \forall x(t), t \in [0, t_f] \tag{2.49}$$

is satisfied.

The same example is treated next but with a cubic terminal constraint. For a given number of terms considered in SSM, we are also interested in analyzing the error of the HJB equation with respect to a set of initial conditions. The problem is as follows;

$$\begin{aligned}&\text{Minimize } J = \frac{1}{2}\int_0^{t_f} u^2 dt \\ &\text{subject to } \quad \dot{x} = u, \quad \psi = x(t_f) + \varepsilon x(t_f)^3 = \psi_f \ ; x(0), \psi_f, \text{and } t_f = \textit{given.}\end{aligned} \tag{2.50}$$

The HJB equation to the above OCP is the same as that of Eq. (2.41) but the boundary condition is,

$$J_x^*(x(t_f), t_f) = \lambda(t_f) = \nu\psi_x(x(t_f)) = \nu(1 + 3\varepsilon x(t_f)^2) \tag{2.51}$$

The SSM is applied with a seventh-order polynomial expansion to approximate the cost-to-go as,

$$\begin{aligned}&J^*(x(t),t) = \{S_0x^2 + S_1x^3 + S_2x^4 + S_3x^5 + S_4x^6 + S_5x^7\} + \\ &\{N_1\nu^2 + N_2\nu^3 + N_3\nu^4 + N_4\nu^5 + N_5\nu^6 + N_6\nu^7\} + \nu\{R_0x + R_1x\nu + R_2x\nu^2 + \\ &+R_3x\nu^3 + R_4x\nu^4 + R_5x\nu^5 + K_0x^2 + K_1x^2\nu + K_2x^2\nu^2 + K_3x^2\nu^3 + K_4x^2\nu^4 + \\ &Q_0x^3 + Q_1x^3\nu + Q_2x^3\nu^2 + Q_3x^3\nu^3 + P_0x^4 + P_1x^4\nu + P_2x^4\nu^2 + \\ &H_0x^5 + H_1x^5\nu + O_1x^6 - \psi_f\}\end{aligned} \tag{2.52}$$

The following system of differential equations for the gains is obtained by substituting Eq. (2.52) into Eq. (2.41) and utilizing Eq.(2.51):

$$\begin{aligned}&\dot{S}_0 = 2S_0^2; S_0(t_f) = 0 \\ &\dot{S}_1 = 6S_1S_0; S_1(t_f) = 0 \\ &\dot{S}_2 = 8S_2S_0 + \frac{9}{2}S_1^2; S_2(t_f) = 0 \\ &\dot{S}_3 = 12S_1S_2 + 10S_3S_0; S_3(t_f) = 0 \\ &\dot{S}_4 = 15S_1S_3 + 12S_4S_0 + 8S_2^2; S_4(t_f) = 0 \\ &\dot{S}_5 = 18S_1S_4 + 20S_2S_3 + 14S_5S_0; S_5(t_f) = 0\end{aligned}$$

$$\dot{N}_1 = \frac{1}{2}R_0^2; N_1(t_f) = 0$$
$$\dot{N}_2 = R_1R_0; N_2(t_f) = 0$$
$$\dot{N}_3 = R_2R_0 + \frac{1}{2}R_1^2; N_3(t_f) = 0$$
$$\dot{N}_4 = R_3R_0 + R_1R_2; N_4(t_f) = 0$$
$$\dot{N}_5 = R_4R_0 + R_1R_3 + \frac{1}{2}R_2^2; N_5(t_f) = 0$$
$$\dot{N}_6 = R_5R_0 + R_1R_4 + R_2R_3; N_6(t_f) = 0$$

$$\dot{R}_0 = 2S_0R_0; R_0(t_f) = 1$$
$$\dot{R}_1 = 2S_0R_1 + 2K_0R_0; R_1(t_f) = 0$$
$$\dot{R}_2 = 2S_0R_2 + 2(K_1R_0 + R_1K_0); R_2(t_f) = 0$$
$$\dot{R}_3 = 2S_0R_3 + 2(K_2R_0 + R_2K_0 + R_1K_1); R_3(t_f) = 0$$
$$\dot{R}_4 = 2S_0R_4 + 2(K_3R_0 + R_3K_0 + R_1K_2 + R_2K_1); R_4(t_f) = 0$$
$$\dot{R}_5 = 2S_0R_5 + 2(K_4R_0 + R_4K_0 + R_1K_3 + R_3K_1 + R_2K_2); R_5(t_f) = 0$$

$$\dot{K}_0 = 4K_0S_0 + 3S_1R_0; K_0(t_f) = 0$$
$$\dot{K}_1 = 4K_1S_0 + 3(S_1R_1 + Q_0R_0) + 2K_0^2; K_1(t_f) = 0$$
$$\dot{K}_2 = 4K_2S_0 + 3(S_1R_2 + Q_1R_0 + Q_0R_1) + 4K_1K_0; K_2(t_f) = 0$$
$$\dot{K}_3 = 4K_3S_0 + 3(S_1R_3 + Q_2R_0 + Q_0R_2 + Q_1R_1) + 4K_2K_0 + 2K_1^2; K_3(t_f) = 0$$
$$\dot{K}_4 = 4K_4S_0 + 3(S_1R_4 + Q_3R_0 + Q_0R_3 + Q_1R_2 + Q_2R_1) + 4(K_1K_2 + K_3K_0);$$
$$K_4(t_f) = 0$$

$$\dot{Q}_0 = 6(Q_0S_0 + K_0S_1) + 4S_2R_0; Q_0(t_f) = \varepsilon$$
$$\dot{Q}_1 = 6(Q_1S_0 + K_1S_1 + K_0Q_0) + 4(S_2R_1 + P_0R_0); Q_1(t_f) = 0$$
$$\dot{Q}_2 = 6(Q_2S_0 + K_2S_1 + K_0Q_1 + K_1Q_0) + 4(S_2R_2 + P_0R_1 + P_1R_0); Q_2(t_f) = 0$$
$$\dot{Q}_3 = 6(Q_3S_0 + K_3S_1 + K_0Q_2 + K_2Q_0 + K_1Q_1) + 4(S_2R_3 + P_0R_2 + P_2R_0 + P_1R_1);$$
$$Q_3(t_f) = 0$$

$$
\begin{aligned}
&\dot{P}_0 = 8(P_0S_0 + S_2K_0) + 5S_3R_0 + 9S_1Q_0; P_0(t_f) = 0\\
&\dot{P}_1 = 8(P_1S_0 + S_2K_1 + P_0K_0) + 5(S_3R_1 + R_0H_0) + 9S_1Q_1 + \frac{9}{2}Q_0^2; P_1(t_f) = 0\\
&\dot{P}_2 = 8(P_2S_0 + S_2K_2 + P_1K_0 + P_0K_1) + 5(S_3R_2 + R_1H_0 + R_0H_1) + 9(S_1Q_2 + Q_0Q_1);\\
&P_2(t_f) = 0\\
&\dot{H}_0 = 12(S_2Q_0 + S_1P_0) + 10(S_0H_0 + K_0S_3) + 6S_4R_0; H_0(t_f) = 0\\
&\dot{H}_1 = 12(S_2Q_1 + S_1P_1 + Q_0P_0) + 10(S_0H_1 + K_1S_3 + K_0H_0) + 6(S_4R_1 + R_0O_1); H_1(t_f) = 0\\
&\dot{O}_1 = 15(S_3Q_0 + S_1H_0) + 12(S_4K_0 + O_1S_0) + 16S_2P_0 + 7S_5R_0; O_1(t_f) = 0
\end{aligned}
\tag{2.53}
$$

The above differential equations can be solved analytically and the following is the series representation for the cost-to-go and feedback control obtained from Eq.(2.52) and Eq.(2.11), respectively:

$$
J^* = \nu\left\{\begin{aligned}
&\frac{153}{40}\varepsilon^2(t-t_f)^5\nu^5 + \frac{117}{8}\varepsilon^2(t-t_f)^4\nu^4x + \frac{81}{4}\varepsilon^2(t-t_f)^4\nu^3x^2 + \frac{9}{2}\varepsilon^2(t-t_f)x^4\nu +\\
&+18\varepsilon^2(t-t_f)^2x^3\nu^2 + \varepsilon(t-t_f)^3\nu^3 + 3\varepsilon(t-t_f)^2\nu^2x + 3\varepsilon(t-t_f)x^2\nu +\\
&+\frac{1}{2}(t-t_f)\nu + (x + \varepsilon x^3 - \psi_f)
\end{aligned}\right\}
\tag{2.54}
$$

$$
u = -\lambda = -\left[\begin{aligned}
&\left(\frac{81}{2}\varepsilon^2(t-t_f)^3\nu^4 + 6\varepsilon(t-t_f)\nu^2\right)x + 3\varepsilon\left(1 + 18\varepsilon^2(t-t_f)^2\nu^2\right)\nu x^2 +\\
&+18\varepsilon^2(t-t_f)\nu^2x^3 + \left(1 + 3\varepsilon(t-t_f)^2\nu^3 + \frac{117}{8}\varepsilon^2(t-t_f)^4\nu^5\right)
\end{aligned}\right]
\tag{2.55}
$$

At every time step, the terminal Lagrange multiplier can be computed from the reversion of the series obtained by using Eqs. (2.15) and(2.54).

The approximation error involved with the methodology to the HJB solution can be evaluated by substituting the J^* obtained in Eq. (2.54) back into the HJB equation. The absolute error that is a function of $x(t)$ and t can be given as,

$$HJB_E(x(t),t) = \left| \left(\frac{\partial J^*(x(t),t)}{\partial t} - (-\min_u (H(x(t),u,\frac{\partial J^*(x(t),t)}{\partial x},t)) \right) \right| \tag{2.56}$$

Depending upon the value of ε and ψ_f, the given terminal constraint can up to have three distinct real roots. For the parameters selected for a numerical solution: $\varepsilon = -0.9$, $\psi_f = 0.1$ and $t_f = 1$, the cubic terminal constraint considered has three distinct real roots, {-1.1009,-1, 0.1009}. A seventh-order expansion is applied with a sixth-order series reversion for evaluating ν. A non-smooth value function is obtained even for this simple example. Figure (2.2) shows the tri-modal value function with three distinct surfaces plotted with respect to the arbitrary initial conditions and the running time. The error in the HJB solution given as Eq. (2.56) is plotted with respect to the set of initial conditions: (-1.5, 1.5). Figure 2.3 shows the error, plotted using a logarithmic scale. The error in the HJB equation also shows a tri-modal pattern. Any initial condition, if it is in the vicinity of the solution of the given terminal constraint, the error in the HJB equation is very less. The SSM also chooses the appropriate surface of the cost-to-go function to satisfy the terminal constraint. It is also noticed that the given initial conditions which satisfy the following condition,

$$\psi_x\big|_{t=t_0} = 0 \tag{2.57}$$

do not lead to the convergence of SSM. This observation is demonstrated in Figs. 2.2 and 2.3, that SSM has no converging solution at the solution of Eq.(2.57), i.e. $x(0) = \pm 0.6086$ or in the neighborhood of these specific points.

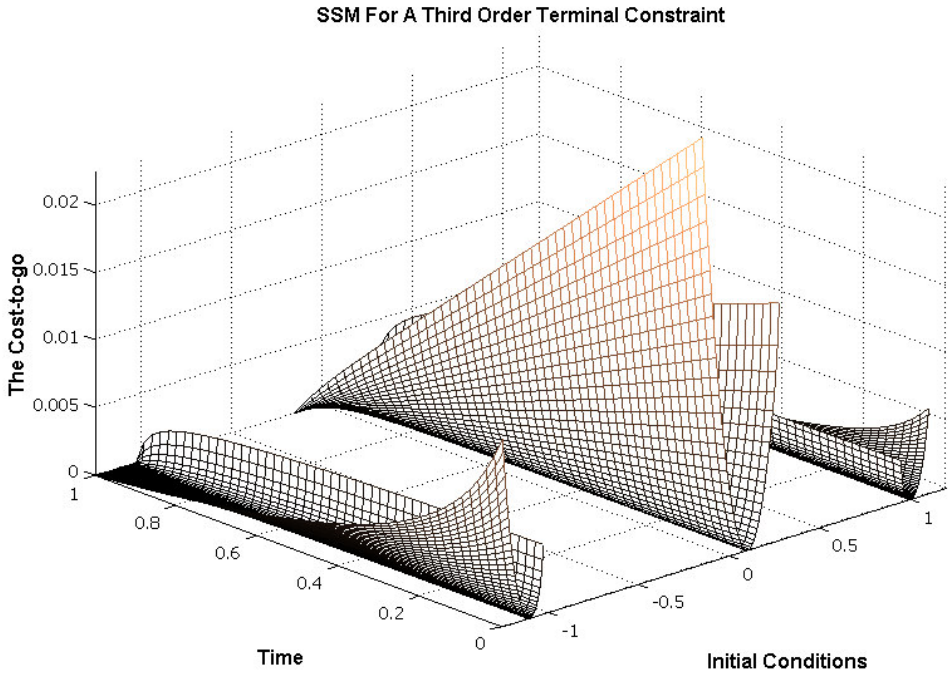


Fig. 2.2 A Tri-modal Cost-to-go Function Obtained by Using SSM

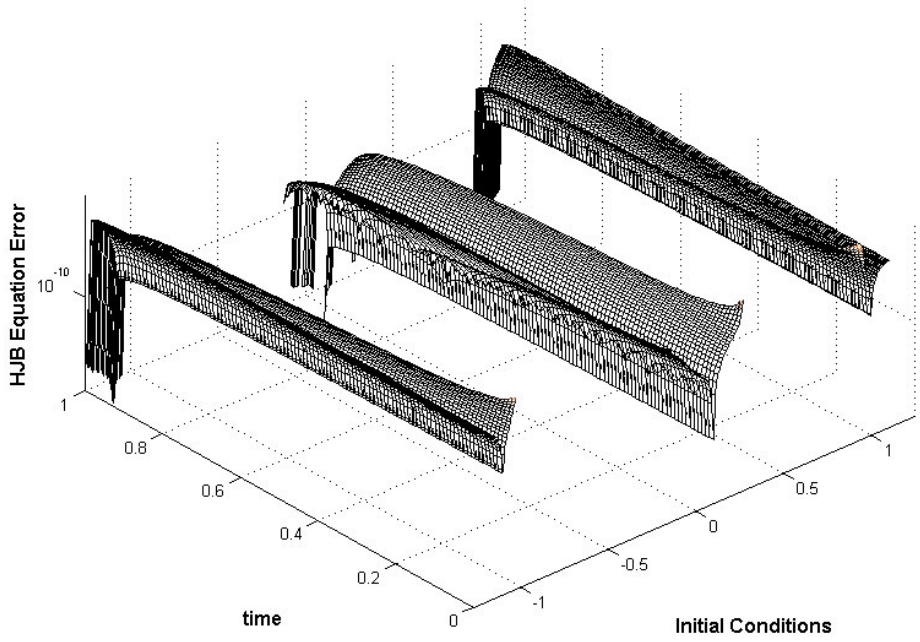


Fig. 2.3 The HJB Equation Error by Using the Seventh-order SSM Approximation

Any initial condition that is outside the domain of convergence of the seventh-order series yields a high approximation error to the HJB equation and the resulting feedback control obtained does not stabilize the system.

It is impossible to globally approximate such discontinuous functions using a single polynomial series. This is a limitation of SSM. However, on the positive side, the Lagrange multiplier-based series solution eliminates the need to solve the terminal constraint equation explicitly. Another advantage of SSM is that it automatically chooses the correct optimal control, appropriate for the initial condition chosen, depending on its proximity to one of the roots of the constraint, when the convergence criterion is satisfied and the required number of terms is included in the series expansion.

2.6 A Numerical Example of a One-dimensional Nonlinear System

In the previous section, the closed-form solution of the HJB equation for a 1-D example was obtained because of the LQ nature of the optimal control problem (OCP). In general, the set of ordinary differential equations cannot be solved analytically and the numerical solution will be required to compute all the gains. The accuracy of the optimal feedback solution will also be dependent on the order of the series utilized. To illustrate these important facts, we consider a one dimensional nonlinear example involving a nonlinear terminal constraint, presented below:

Minimize:

$$J = \frac{1}{2}\int_{t_0}^{t_f} u^2 dt$$

subject to

$$\begin{aligned} &\dot{x} = -x - 0.5x^2 + u; \\ &x(t_0) = x_0 = 0.1; t_0 = 0; t_f = 5 \\ &\psi(x(t_f)) - \psi_f \equiv x(t_f) + 0.3x(t_f)^2 - 0.5 = 0 \end{aligned} \tag{2.58}$$

The HJB equation for the above OCP is,

$$\frac{\partial J^*}{\partial t} = \frac{1}{2}\left(\frac{\partial J^*}{\partial x}\right)^2 - \frac{\partial J^*}{\partial x}(-x - 0.5x^2); J^*(x(t_f)) = 0 \tag{2.59}$$

whereas $x(t_f)$ satisfies, ψ.

A 4th order series expansion of the cost-to-go in x and ν is considered for solving Eq. (2.59). As described in Sec. 2.5, substituting Eq. (2.38) of J^* into Eq. (2.59)

and collecting the coefficients of various monomials of x and ν will result in the required differential equations for the gains. Transversality conditions for optimality can be utilized to find the boundary conditions for integrating these gain differential equations backward in time, resulting in gain functions that can be stored a priori. The closed-loop feedback control law requires the calculation of ν at each time step, which can be achieved by utilizing the method of series reversion.

Figure 2.4 demonstrates the numerical results of the application of the higher-order feedback solution. The feedback solution (solid line) is compared with the open-loop solution, shown using a black dotted line. It can be clearly seen that the third-order feedback solution is indistinguishable from that of the open-loop solution, whereas the first and second order solutions show significant errors in the midcourse and the terminal regions.

Figure 2.5 demonstrates that the given OCP has multiple open-loop solutions depending upon the initial guess utilized. Two different open-loop solutions are plotted along with the linear, second, and third-order feedback solutions for one initial condition. It is evident that the series solution method provides a structured mechanism to solve the given HJB equation yielding the minimum solution which satisfies both the necessary and sufficient conditions for optimality and the terminal constraint, exactly. As further evidence of the utility of SSM, the HJB error defined as Eq. (2.56) is plotted in Fig. 2.6 for a wide range of initial conditions. It is demonstrated that all initial conditions chosen between [0, 1] converges to the appropriate terminal solution with a terminal error of order 1e-10 (numerically zero).

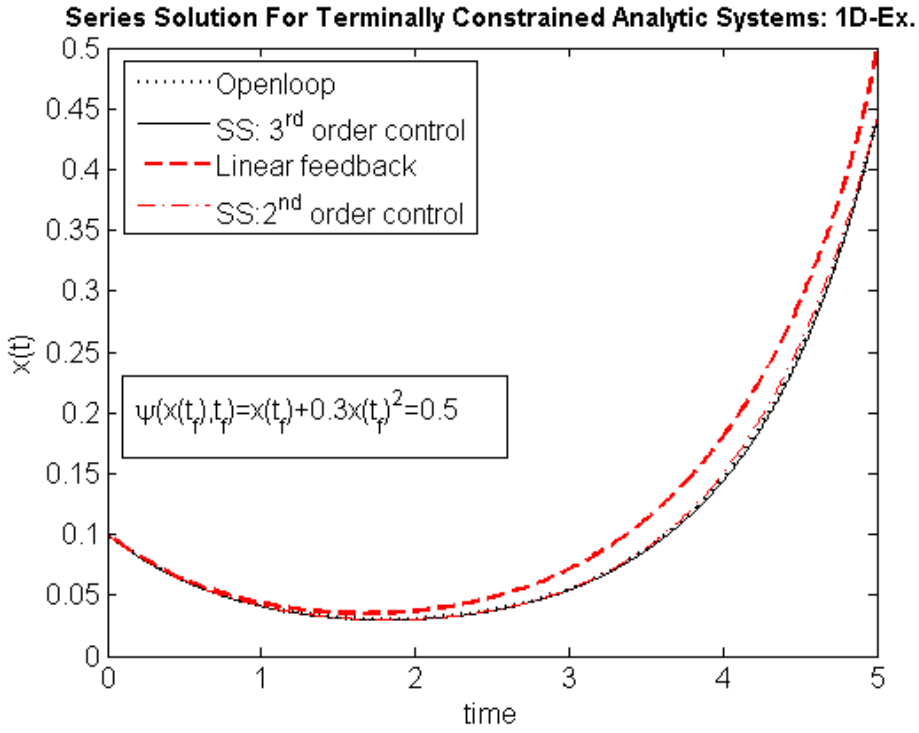


Fig. 2.4 Series Solution Method for Terminally Constrained Nonlinear System

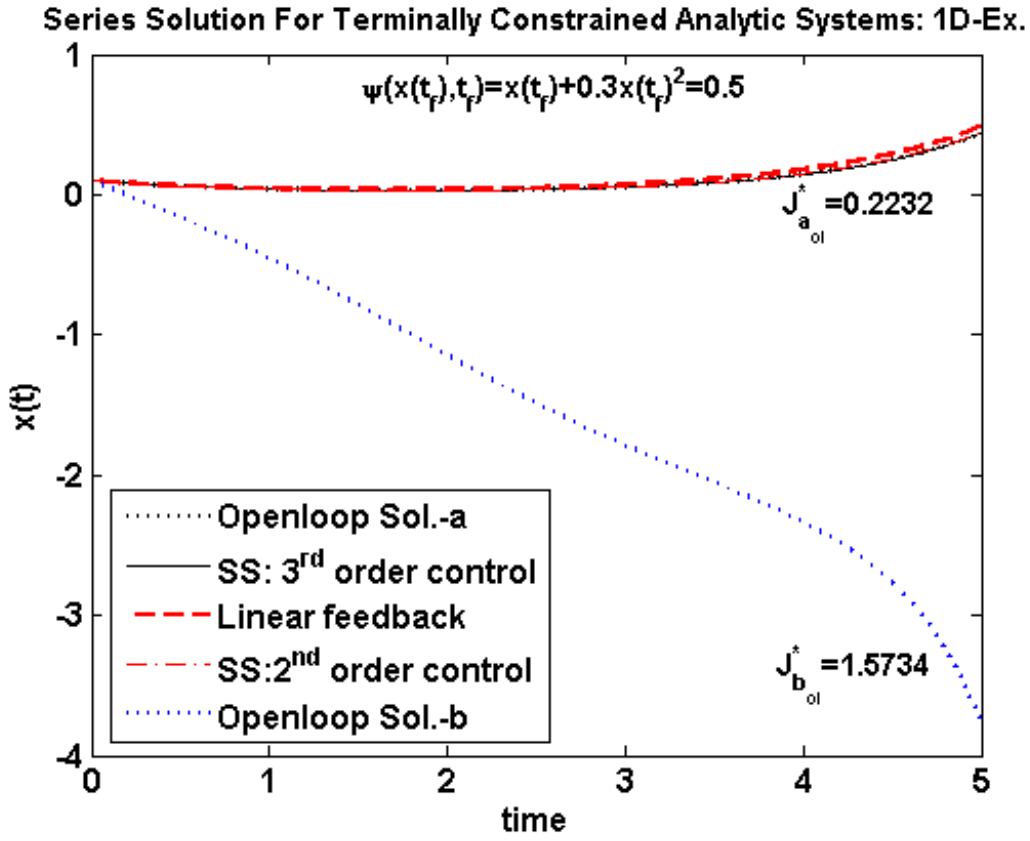


Fig. 2.5 Multiple Open-loop Solutions for 1-D Terminally Constrained Nonlinear System

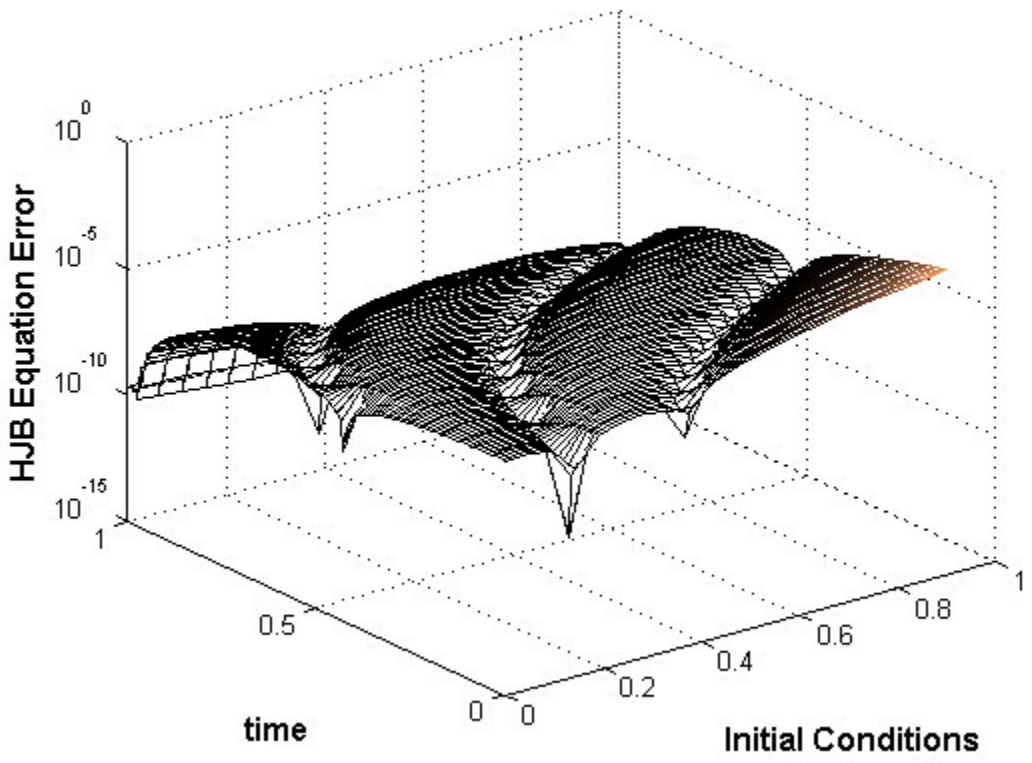


Fig. 2.6 The HJB Error with Respect to Initial Conditions for 1-D Nonlinear Example

2.7 Galerkin Approximation Techniques for Optimal Control Problems

Another polynomial expansion method, significantly different from SSM, is the Galerkin Approximation Technique (GAT). A version of this technique was first introduced by Beard and Saridis[27] for OCPs without terminal constraints. They developed a technique for obtaining an approximate solution of the HJB equation via a linear partial differential equation defined as the generalized Hamilton-Jacobi-Bellman (GHJB) equation. Theirs is an iterative method which requires a stabilizing feedback

control at the first step for the convergence of subsequent policy iterations to the optimal feedback solution. In this section, we present a direct non-iterative algorithm to solve the time-dependent HJB equation by employing the Galerkin's techniques as used in Ref. [28].The major motivation to present this new algorithm is to utilize the spectral properties in series formalism with Galerkin applications. The methodology, especially, becomes attractive because of its flexibility to a variety of bases functions in the formulation of the cost-to-go expansion. Moreover, it is not limited to the polynomial structure of the performance index and the dynamical system as well. To lay out the complete solution procedure, a 1-D problem is addressed with the following steps,

1. Assume the series of the optimal return function as,

$$J^*(x,t) = \sum_{i=0}^{N} c_i(t) T_i(x) \tag{2.60}$$

where c_i are time dependent unknown gain coefficients and T_i are orthogonal polynomial bases functions. N is the number of terms considered in the expansion of J^*.

2. Next, to obtain the gain coefficients, c_i substitute Eq. (2.60) into Eq. (2.9) that produces the residual, E in the HJB equation due to the truncated expansion of J^*,

$$E(x(t),t) \equiv \left(\frac{\partial (c_i T_i)}{\partial t} + \underbrace{\min}_{u} \left(H(x(t),u(t), \frac{\partial (c_i T_i)}{\partial x}, t \right) \right) \tag{2.61}$$

Taking into account that a better approximation can be obtained by making the residuals smaller, the orthogonality between residuals and the weight functions can be enforced on the span of the basis function T_i with its appropriate weight terms, *w*. Galerkin spectral techniques are imposed to yield the required number of gain differential equations as shown below:

$$\int_{\Omega_x} T_j w E dx \equiv \{\dot{c}_j - f(c_j)\} = 0 \tag{2.62}$$

In the specified spatial domain Ω_x, the integrals formulated in Eq. (2.62) can be computed symbolically.

3. Now, to perform the numerical integration of the system equations (2.62) backward in time, the boundary conditions are obtained by projecting the weighted residual obtained from the transversality condition or Eq. (2.10) on the span of the applied bases functions.

4. Substitute the stored gains into Eq. (2.7) and Eq. (2.11) to construct the optimal feedback solution.

To explain these steps clearly, we begin with a simple 1-D LQ example given as,

$$\begin{aligned}
&\text{Min } J = \int_{t_0}^{t_f} (x^2 + \|u\|_R^2) dt \\
&\text{subject to} \\
&\dot{x} = u; \\
&x(t_0)=0.5 \text{ , } t_0 = 0; t_f = 1 \\
&R = 1
\end{aligned} \tag{2.63}$$

This OCP has no hard or soft constraint at the final time.

To solve for the HJB equation, the cost-to-go expansion can be taken as,

$$J^*(x,t) = c_0(t)T_0(x) + c_1(t)T_1(x) + c_2(t)T_2(x) \tag{2.64}$$

where $T_i(x), i = 0,1,2$ are the Chebyshev polynomials of the first kind in $\Omega_x = [-1,1]$ (Appendix B).

Following the approach described above, the symbolic form of gain differential equations with boundary conditions are obtained as,

$$\begin{aligned}
&\dot{c}_0 = 2c_2^2 + \frac{c_1^2}{4} - \frac{1}{2}; c_0(1) = 0, \\
&\dot{c}_1 = 2c_1c_2; c_1(1) = 0, \\
&\dot{c}_2 = 2c_2^2 - \frac{1}{2}; c_2(1) = 0,
\end{aligned} \tag{2.65}$$

It can be clearly noticed that unlike what we observe in SSM which produces a Riccati equation and a system of linear coupled gain differential equations; the GAT produces a set of coupled and nonlinear gain differential equations. These equations can be integrated numerically and the stored gains can be utilized to construct the HJB solution.

For the LQ problem given as Eq.(2.63), Fig. 2.7 shows the absolute difference in the state trajectories yielded by comparing the open-loop solution with the feedback

solutions. It is shown that there is no difference between the numerical simulations performed by using SSM and the GAT. In the next section, nonlinear examples are considered to understand the importance of each methodology.

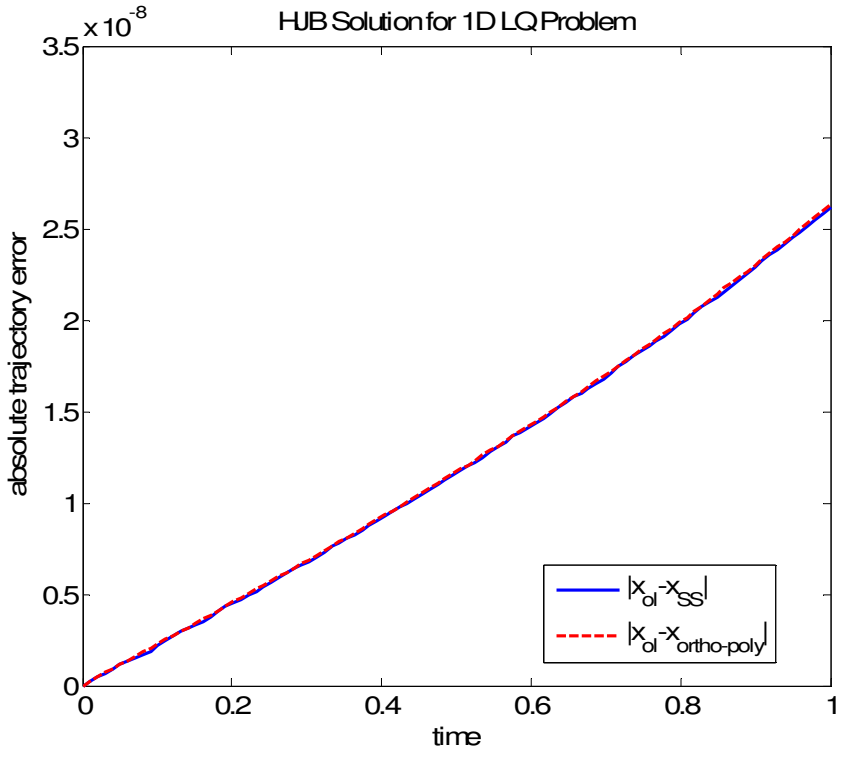


Fig. 2.7 Comparison of the Series Solution and Galerkin Technique with Respect to Open-loop Solution

For problems with terminal constraints, each terminal Lagrange multiplier can be treated as an additional spatial dimension in the extended state-space in the appropriate domain Ω_ν. Usually, Ω_ν cannot be fixed a priori, however, it can be the same as Ω_x with

proper scaling of the variables and time. The general representation of optimal return function can be given in $x \in \Re^n$ and $\nu \in \Re^p$ as,

$$J^*(x,\nu,t) = \sum_{i=0}^{N}\sum_{j=0}^{M}\sum_{k=0}^{P}\ldots\sum_{i_1=0}^{N_1}\sum_{j_1=0}^{M_1}\sum_{k_1=0}^{P_1}\ldots c_{ij\ldots ki_1j_1\ldots k_1}(t)\{T_i(x_1)T_j(x_2)\ldots T_k(x_n)\}\{T_{i_1}(\nu_1)T_{j_1}(\nu_2)\ldots T_{k_1}(\nu_p)\} \tag{2.66}$$

where the order of polynomial of the cost-to-go, *K* can be given as,

$$K = N + M + \ldots + P + N_1 + M_1 + \ldots + P_1 \tag{2.67}$$

Following the steps as explained above, the gain differential equations and the boundary conditions can be obtained as,

$$\begin{aligned}&\int_{\Omega_{\nu_p}}\ldots\int_{\Omega_{\nu_1}}\int_{\Omega_{x_n}}\ldots\int_{\Omega_{x_1}} T_i(x_1)T_j(x_2)\ldots T_k(x_n)\}W(x,\nu)E\{T_{i_1}(\nu_1)T_{j_1}(\nu_2)\ldots T_{k_1}(\nu_p)\}dx_1\ldots dx_n d\nu_1\ldots d\nu_p \\ &\equiv\{\dot{c}_{ij\ldots ki_1j_1\ldots k_1} - f(c_{ij\ldots ki_1j_1\ldots k_1})\} = 0 \\ &\text{where, } W(x,\nu) = w(x_1)\ldots w(x_n)w(\nu_1)\ldots w(\nu_p)\end{aligned} \tag{2.68}$$

$$c_{ij\ldots ki_1j_1\ldots k_1}(t_f) = \int_{\Omega_{\nu_p}}\ldots\int_{\Omega_{\nu_1}}\int_{\Omega_{x_n}}\ldots\int_{\Omega_{x_1}}\begin{pmatrix} T_i(x_1)T_j(x_2)\ldots T_k(x_n)\}W(x,\nu)\{J^*(x(t_f))+\ldots \\ \ldots\nu^T(\psi(x(t_f))-\psi_f)\}\{T_{i_1}(\nu_1)T_{j_1}(\nu_2)\ldots T_{k_1}(\nu_p)\}\end{pmatrix} dx_1\ldots dx_n d\nu_1\ldots d\nu_p \tag{2.69}$$

While integrating the closed-loop dynamics of terminally constrained problem, the feedback form requires the terminal Lagrange multiplier, which can be computed by using the series reversion after substituting Eq. (2.66) into Eq. (2.15).

2.8 Error Analysis with Respect to the Feedback Methods

Here, we attempt to compare and contrast the performances of SSM and GAT.

Based on the power series approximations, two optimal feedback methods are established to solve the OCPs. Each method has its own applicability depending on the nature of the plant dynamics, cost function and the involved nonlinearity. The error structure associated with each method depends on the truncation of power series utilized for the cost-to-go and the series reversion. To analyze the efficacy of each feedback method, the following cases of 1-D nonlinear examples with a quadratic cost are considered as,

CASE A: Nonlinear Third-order 1-D System

$$\begin{aligned} &\text{Min } J = \int_{t_0}^{t_f} (x^2 + \|u\|_R^2) dt \\ &\text{subject to} \\ &\dot{x} = -x - 2x^2 - 0.5x^3 + u; x(t_0) = 0.5, \\ &R = 1; \quad t_0 = 0; t_f = 1 \end{aligned} \tag{2.70}$$

CASE B: Nonlinear OCP but No Linear Term in the Dynamical System

$$\begin{aligned} &\text{Min } J = \int_{t_0}^{t_f} (x^2 + \|u\|_R^2) dt \\ &\text{subject to} \\ &\dot{x} = -2x^2 - 4x^3 + u; x(t_0) = 0.5, \\ &R = 1; t_0 = 0; t_f = 1 \end{aligned} \tag{2.71}$$

In case A, the plant dynamics has relatively high nonlinearity in the second order term compared to the first and third order terms; case B is chosen without any linear term in the dynamical system. Usually such cases require more number of terms to

approximate the cost-to-go expansion. For both the cases A& B, sixth order polynomials are applied to use SSM for solving the respective HJB equations. The Chebyshev polynomials are also considered up to the sixth-order for showing the comparison between SSM and GAT. As an output, the absolute error between open-loop trajectory and the feedback solutions are plotted to emphasize the numerical accuracy with each method. To present the comparative study of the cost-to-go results for these examples, Table 2.1 is also constructed as,

Table 2.1 The Cost-to-go Analysis

	J^* Open-loop	J^* Series Solution (6^{th} Order)	J^* Galerkin Method (6^{th} Order)
CASE A	0.06535320310710	0.06535322438679	0.06535337797984
CASE B	0.08569804986542	0.08680463083558	0.08571043092278

For case A, Fig. 2.8 clearly shows the dominance of SSM over GAT which yields more error than that obtained by using SSM. In contrast, the results of case B lead to a different conclusion. Within the same level of truncation, as shown in Fig. 2.9, GAT produces less error in the state trajectory; however, this problem-oriented comparison is conclusive for the given initial condition only.

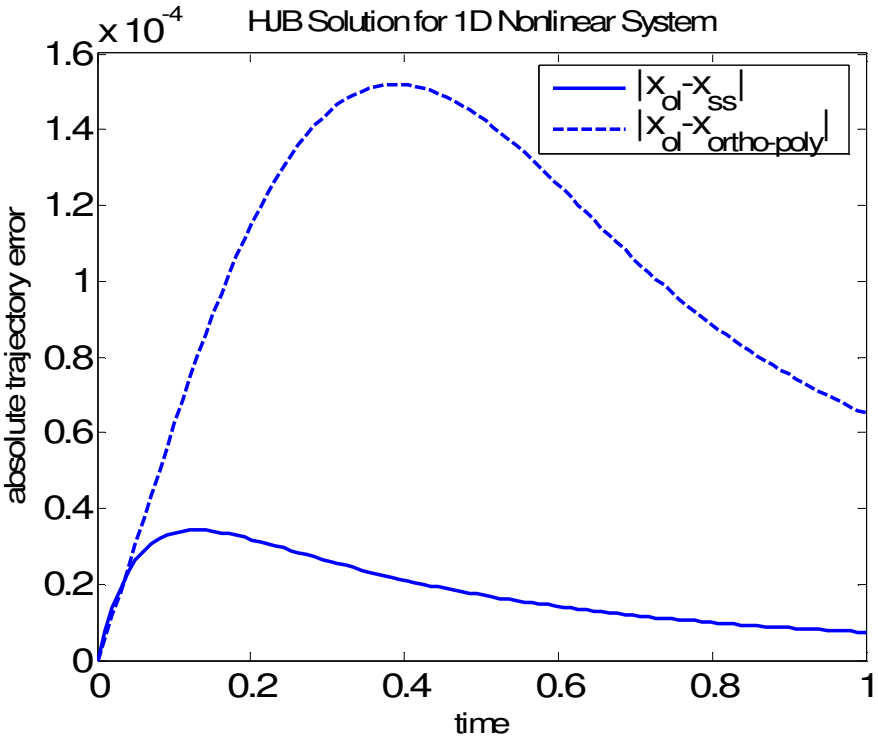


Fig. 2.8 CASE A: Comparison of the Series Solution and Galerkin Technique with Respect to the Open-loop Solution for Highly Nonlinear System

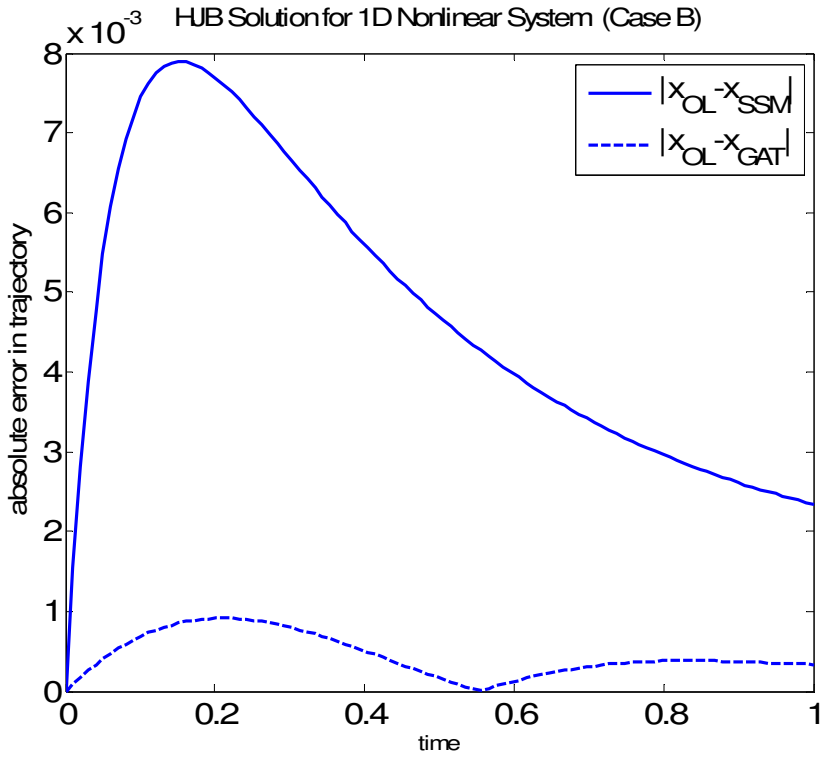


Fig. 2.9 CASE B: Comparison of SSM and GAT with Respect to the Open-loop Solution for Highly Nonlinear System

To present a general comparison of series approximation, the cost-to-go and the error in the HJB solution for each method defined as Eq. (2.56) and Eq. (2.61) , respectively can be plotted with respect to Ω_{x_0} defined as $\{x_0 = x(t_0) \big| x(t) \in \Omega_x\}$.

In $\Omega_{x_0} = (-1,1)$, Fig. 2.10 shows that the cost-to-go obtained by using SSM and GAT are almost the same. For further analysis, the error structure of the HJB solution is demonstrated in Fig. 2.11 for both the methods. It clearly shows that the error produced by using the GAT is approximately uniform for a wider range of initial conditions whereas that obtained by using SSM is much less locally. Depending on the order of

truncation, the error plot also shows the region of convergence associated with each method. Except the neighborhood of the endpoints of Ω_{x_0}, the application of SSM is more advantageous than the use of the GAT.

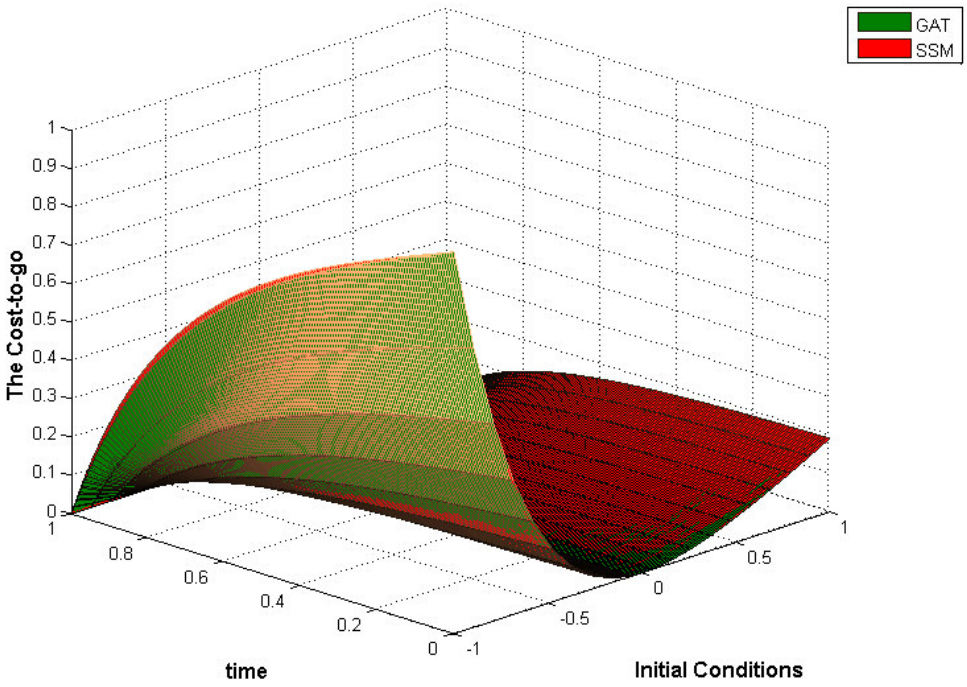


Fig. 2.10 CASE A: The Cost-to-go with Respect to Initial Conditions

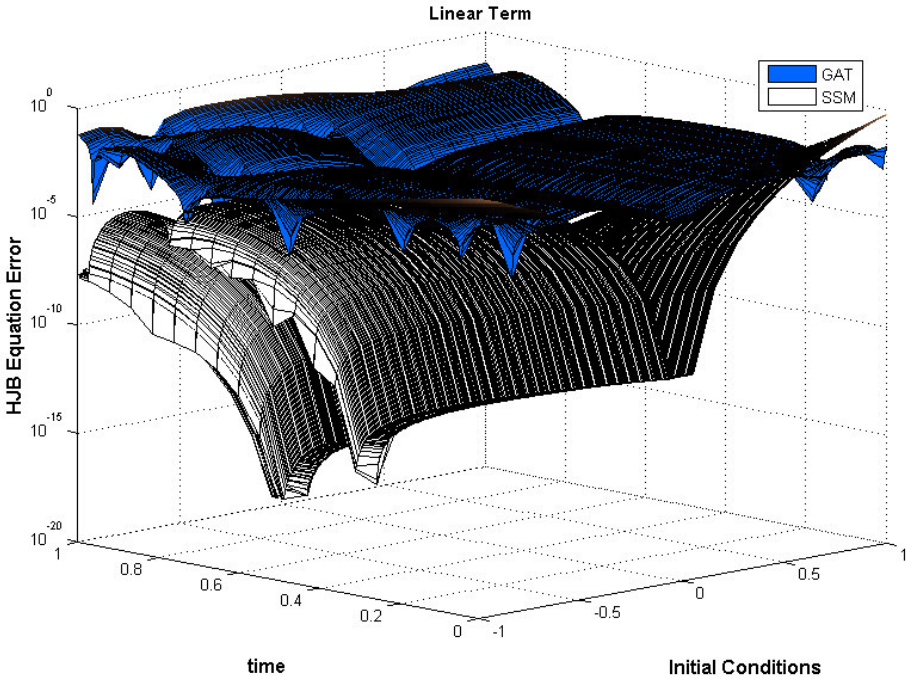


Fig. 2.11 CASE A: The HJB Error with Respect to Initial Conditions

In general, SSM requires the presence of linear terms in the dynamical system. Case B does not contain the linear term in the OCP. Hence, many of the gains associated with quadratic, cubic or higher order terms in x and v considered in the cost-to-go expansion are eliminated. The GAT provides an advantage for such problems and provides the appropriate approximation with a smaller number of gains. An OCP for a simple bilinear system with an analytical solution is treated as an example to clarify the previous statements:

Minimize:

$$J = \frac{1}{2}\int_{t_0}^{t_f} u^2 dt$$
subject to (2.72)
$\dot{x} = xu; x(t_0) = x_0;$
and a point terminal constraint is given as $x(t_f) = x_f$

The above OCP does not contain a linear term. The exact feedback solutions for the state and control are

$$x = x_0 \exp\left(-\frac{t}{t_f}\ln\left(\frac{x_0}{x_f}\right)\right) \tag{2.73}$$

$$u = -\frac{1}{(t_f - t)}\ln\left(\frac{x}{x_f}\right) \tag{2.74}$$

For this example, SSM will require a polynomial expansion which can clearly reflect the logarithmic nature of the exact value function. The SSM can provide an accurate solution within its domain of convergence for this class of problems, only over a small time domain. However, in general, over larger time domains, the standard polynomial expansion will require a large number of higher-order terms leading to a higher computational burden and, it may also not converge.

For the above example, the cost-to-go expansion is carried up to sixth order for both the methods. The cost-to-go and the error in the HJB solutions obtained from each method are also plotted with respect to a set of arbitrary initial condition in Figs. 2.12 and 2.13, respectively.

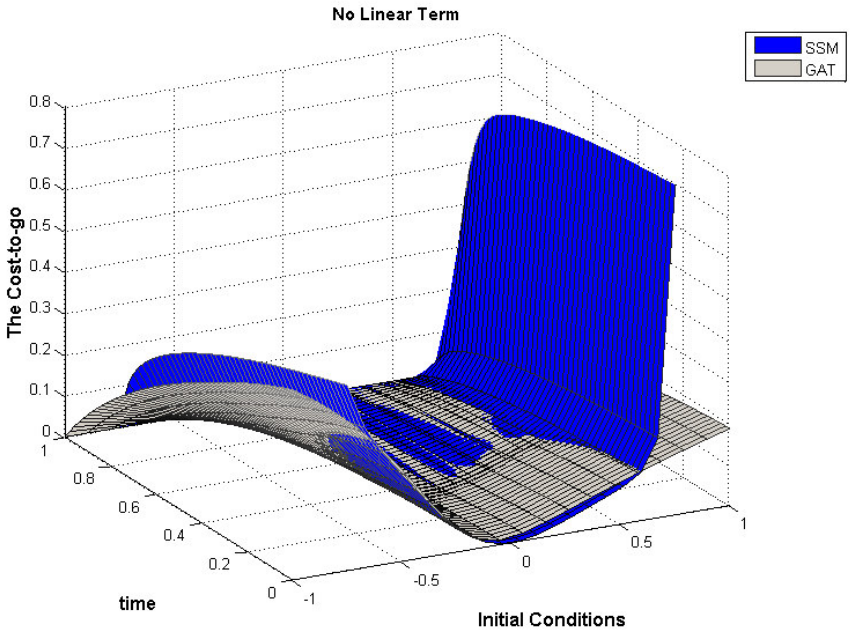


Fig. 2.12 CASE B: The Cost-to-go with Respect to Initial Conditions

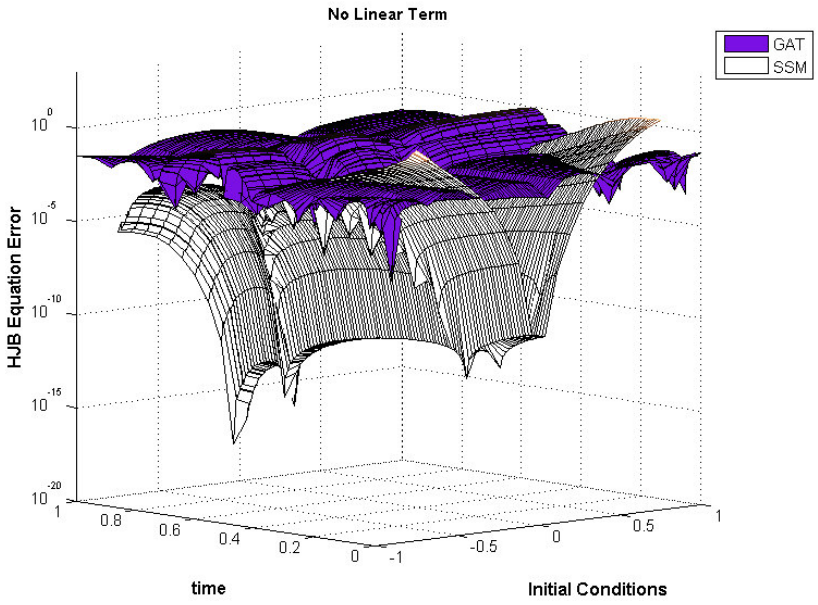


Fig. 2.13 CASE B: The HJB Error with Respect to Initial Conditions

The SSM does not produce a very good approximation of the HJB solution for all the initial conditions considered in $[-1,1]$, because of using the polynomial bases functions, whereas GAT produces a uniform error in the solution. However, for a set of initial conditions in Ω_x, the SSM dominates over GAT in terms of providing a better local convergence. The set of initial condition for which SSM is better than GAT depends on the nonlinearity of the given OCP and the number of terms considered in each method. However, it can be determined within the level of the error tolerance after solving the OCP and substituting the solution into the HJB equation.

As observed for cases A and B, unlike the non-uniform structure of the equation error produced by SSM, GAT allows for a uniform distribution of the approximation error over the domain of the convergence. Even for open-loop stable initial conditions with respect to the origin, the HJB solution from GAT exhibits a higher level of global error than that obtained by using SSM. However, GAT has the ability to accommodate general nonlinearities for a wider range of initial conditions; it utilizes the pseudospectral characteristics of the orthogonal polynomials, producing a fully populated gain vector at each stored step, whereas the SSM only yields a sparse structure of these gains as a result of using polynomial bases functions.

For terminal constraint problems which are the main focus in this development, a general problem-oriented investigation reveals that there is not much numerical difference in the cost-to-go results obtained by using both the feedback methods, if the plant is stable and weakly nonlinear. The feedback method, even for moderate

nonlinearity, can be chosen based on the coefficients of the performance index and the presence of linear terms in the plant dynamics. However if the plant is locally unstable, it is found that SSM dominates GAT in terms of appropriate feedback approximations within the chosen order of truncation. The following nonlinear example with an unstable plant is considered to show the functioning of GAT for a nonlinear terminal constraint problem and a comparison of both the methods is also provided,

Minimize:

$$
\begin{aligned}
&J = \int_{t_0}^{t_f} (x^2 + \|u\|_R^2) dt \\
&\text{subject to} \\
&\dot{x} = x - 0.8x^3 + u; x(t_0) = 0.5 \\
&R = 1; t_0 = 0; t_f = 1; \\
&\psi(x(t_f)) - \psi_f = 0 \equiv x(t_f) + 0.5x(t_f)^2 - 0.25
\end{aligned}
\tag{2.75}
$$

To solve the above problem by using GAT, the cost-to-go expansion is taken up to 8th order in Chebyshev polynomials involving x and ν. The gain equations with the boundary conditions are obtained by using Eqs. (2.68)-(2.69). After storing these gains in the given time-horizon, the closed-loop dynamics is integrated with terminal Lagrange multiplier updated at every instant. To compute the value of ν, the 6th order series reversion process is utilized for solving Eq. (2.15). The given terminal constraint in this problem has two solutions; just like SSM, the Galerkin method also selects the best solution without the necessity of solving the terminal constraint, explicitly. In Fig. 2.14, the solution obtained by using GAT is plotted with the blue dashed line whereas that for the 8th order SSM is presented with the blue bold line. Both methods converge to the

exact solution of the terminal constraint but the Galerkin method exhibits more error in the midcourse trajectory. Computationally, each feedback method requires the same number of gains, if carried up to the same order. Increasing the number of terms or orthogonal polynomials in the cost-to-go expansion improves the feedback solutions but the solution obtained by using SSM shifts closer to the open-loop solution, which is also plotted in the same figure with the black dashed line. The approximation error attained by using each feedback method is compared in Fig. 2.15 for a range of stabilizing initial conditions, $\Omega_{x_0} = [-0.1, 0.6]$. Acceptable terminal error is obtained by using both the feedback methods; however, the performance of the SSM is better than that of the GAT for the unstable systems.

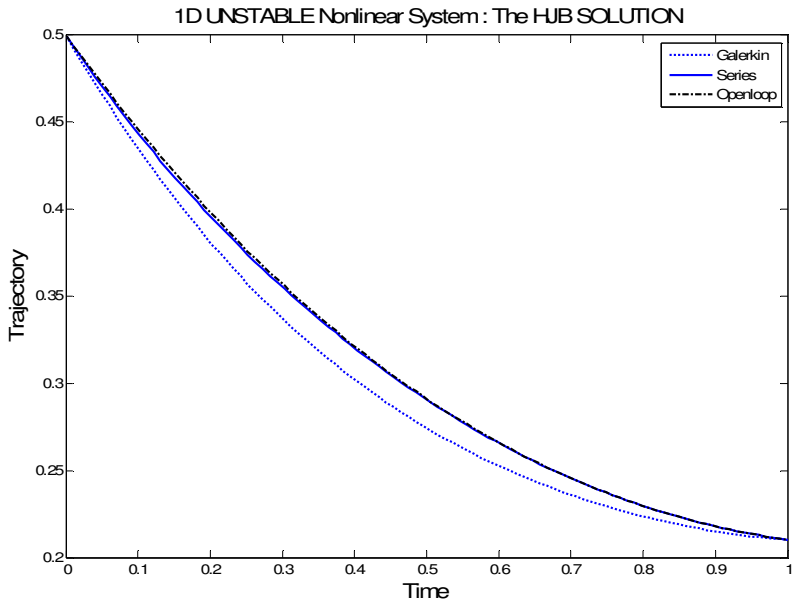


Fig. 2.14 Comparison between the Series Solution and Galerkin Technique with Respect to the Open-loop Solution for 1-D Unstable Nonlinear System

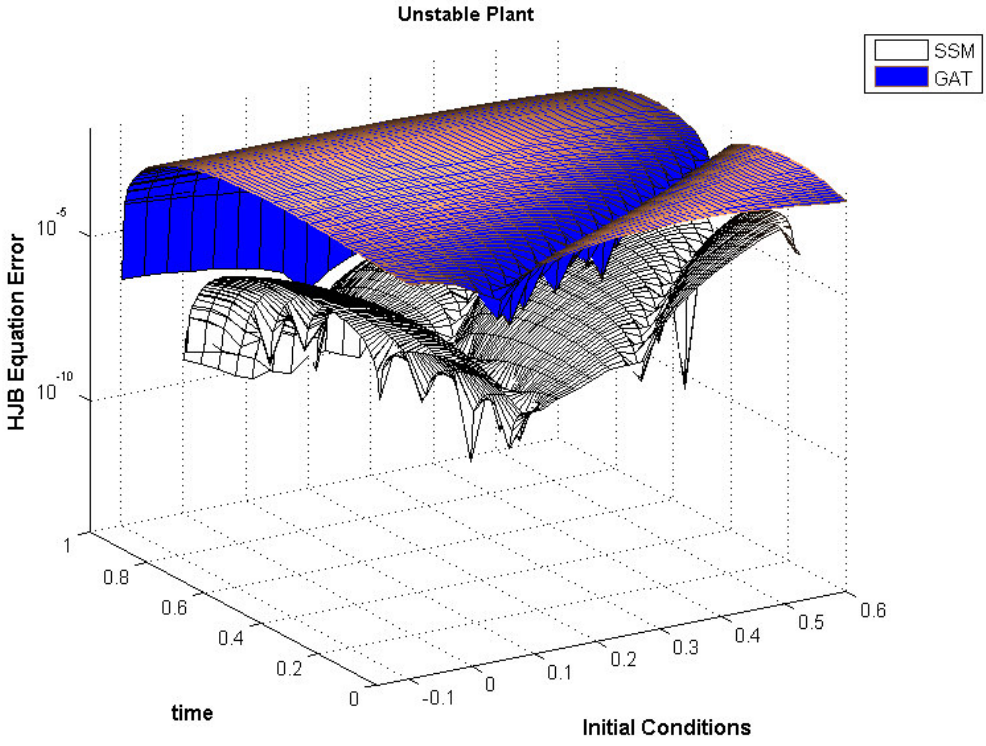


Fig. 2.15 Unstable Third-order Plant: The HJB Error with Respect to Initial Conditions

As required for carrying out the process of collecting the coefficients of the like terms in SSM, GAT also requires a large computational memory to perform the symbolic computation of the integrals formed on the weighted residuals. Especially, in higher dimensional system, it is highly cumbersome to evaluate these multidimensional integrals for real-time processing. The uniqueness of these gains cannot be guaranteed due to the coupled system of nonlinear differential equations. Within the specified tolerance of the error, the state space domain corresponding to a positive definite cost is defined as a region of stability. This region can be estimated by obtaining the intersection of the set of initial conditions applied in receiving the specified error in the HJB solution and a positive definite cost.

General recommendations regarding the suitability of the two methods can be made based upon 1) the presence/lack of linear terms in the dynamical system and 2) the stability of the open-loop system. The SSM presented above is especially attractive when the dimension of ν is much smaller than that of x, since the number of gains is reduced. For weakly nonlinear stable plants or if the nonlinearity diminishes with the higher order terms in the plant dynamics, SSM can be utilized to produce fast, converging results with desired terminal accuracy. On the other hand, when the dynamical system does not have linear terms or the performance index is devoid of quadratic terms, GAT is preferable, due to its wider domain of convergence. This fact is evident from the presented numerical examples. Further more, for the class of unstable plants considered, the SSM provides a better approximation of the cost-to-go, resulting

in a stabilizing feedback law and an HJB solution which is very close to the open-loop solution to the posed OCP.

Further work in this dissertation is continued with the SSM because of the following reasons,

1. For open-loop stable systems, in general, GAT produces higher approximation error than that obtained by using the SSM.
2. Computing the integrals required for GAT is very tedious for higher dimensional systems.
3. Both SSM and GAT require problem-oriented symbolic/numerical computations. Hence, only one of the methods is selected for subsequent analysis.

In the following chapters, SSM applications to solve the governing HJB/HJI equations for higher dimensional dynamical systems are investigated on several examples of nonlinear feedback design for aerospace optimization problems.

CHAPTER III

NUMERICAL IMPLEMENTATIONS

In this chapter we present some examples that illustrate the Polynomial Series Expansion Methodology, particularly, the Series Solution Method (SSM).

Recall that as part of SSM, the cost-to-go function is approximated by using a power series expansion in the system states and terminal Lagrange multipliers. The coefficients in the polynomial expansion represent time-dependent feedback control gains, and these gains are governed by the Riccati equation and uncoupled linear differential equations. The SSM is suited for systems governed by polynomial nonlinearities.

The first example is a two-state, nonlinear problem subject to a single control input. The final time is specified, at which, the final states must lie on some nonlinear surface. This example will show that the solution accuracy improves as more terms are included in the polynomial approximation of the cost-to-go function. The results are compared with a computed open-loop solution. Furthermore, this example will show that the computed feedback solution works well for off-nominal initial conditions, and it will show a key point of the general dynamic programming approach, which is that when a solution to the HJB is computed, the solution is the locally minimizing solution (This contrasts against shooting methods to obtain open-loop solutions, which may only be extremals).

The next example is a rigid body detumbling maneuver. This example will show that although the theoretical terminal Lagrange multipliers are constant, the computed Lagrange multipliers using SSM are only nearly constant. This is due to the inaccuracy of the series reversion, which was briefly discussed in chapter II, section 2.4.

The final example presented is an orbit transfer problem. This nonlinear problem is not governed by polynomial nonlinearities, and so SSM is not directly applicable. Nevertheless, by performing a Taylor series expansion of relevant equations around a trivial nominal trajectory, one can still use SSM. Error analysis is also performed with respect to the order of feedback control by using the open-loop solutions to the respective problems for comparison. It is shown that for small perturbations, highly accurate guidance solutions can be achieved by using a third-order feedback control law.

3.1 A Two-dimensional Example

The following, two-dimensional, example is presented to further illustrate the application of SSM. Two cases are considered below by choosing different sets of parameters for the model. Case A is a highly nonlinear stable plant whereas case B is constructed to check the applicability of SSM for synthesizing the optimal feedback control of an unstable nonlinear dynamical system.

Equations(3.1),(3.2) and(3.3), respectively, show the dynamical model, constraint, and performance index:

$$\ddot{x} = -(m_1 x + m_2 x^2 + m_3 x^3) + Bu \tag{3.1}$$

$$\psi = [\varepsilon_1 x(t_f) + \varepsilon_2 \dot{x}(t_f) + \varepsilon_3 x(t_f)^2 + \varepsilon_4 \dot{x}(t_f)^2 + \varepsilon_5 x(t_f)\dot{x}(t_f)] = \psi_f \tag{3.2}$$

$$J = \int_{t_0}^{t_f} (Q_{11} x^2 + Q_{22} \dot{x}^2 + Ru^2) dt \tag{3.3}$$

Case A: Highly Nonlinear System and Nonlinear Terminal Constraint

This example is parameterized by the following choices for the constants in Eqs. ((3.1)-(3.3)):

$$\begin{aligned} &m_1 = 1; m_2 = 0.8; m_3 = 0.75 \\ &\varepsilon_1 = 0.9; \varepsilon_2 = 0.6; \varepsilon_3 = 0.6; \varepsilon_4 = 0.3; \varepsilon_5 = 0.1; \\ &Q_{11} = Q_{22} = B = R = 1; \\ &t_0 = 0; t_f = 3; x(t_o) = [0.1; 0.2] \\ &\psi_f = 0.5 \end{aligned} \tag{3.4}$$

Unlike the previous one-dimensional example discussed in chapter II, the gains are solved numerically for this example. In the actual implementation of the feedback control law, the gain equations were integrated forward in time, along with the state equations, using the known initial conditions. Although the computational burden increases with such an implementation, storage of the time-varying gains is avoided. In all the examples presented here, the value function is expanded to fourth order and the series reversion process to determine ν is also implemented to fourth order. Furthermore, since ν cannot be determined at the final time, the simulation is stopped just prior to reaching the final time. For each example, the corresponding open-loop optimal solution obtained by using a shooting method is also presented.

Figure 3.1 shows the open-loop and feedback control histories for the nominal and perturbed initial conditions. It is interesting to note that the feedback control histories track their open-loop counterparts closely. Deviations between the open-loop and feedback control histories are evident due to the approximations introduced by the series expansion and the series reversion process. The results for the perturbed initial conditions were obtained by utilizing forward integration only, since the initial conditions of the gains were known from the computations for the nominal problem. The error in satisfying the terminal constraint is approximately 4.8×10^{-5}, for a 10% change in the initial state from its nominal value. The simulation is stopped slightly before reaching the final time of 3 sec because of the singularity at the final time.

Figure 3.2 shows the open-loop and feedback phase portraits for various initial conditions. Even though there is some deviation in the mid-course, the terminal constraint is satisfied quite accurately. Figure 3.3 shows the convergence of the first state toward its optimal solution, as more terms in the series are included. A third-order series solution for the control is seen to be much better than the lower order control approximations.

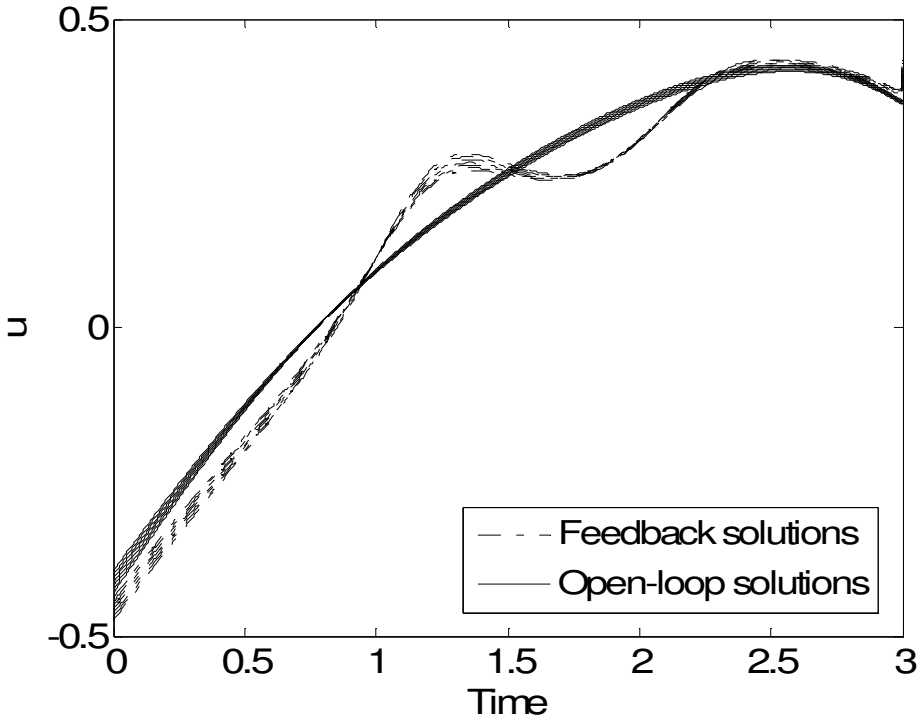


Fig.3.1 Case A: Open-loop and Feedback Control Histories

Another interesting observation is the existence of multiple open-loop solutions for the same initial conditions, depending on the initial guesses for the costates. Since the shooting method utilizes necessary conditions only, the result cannot be guaranteed to be a local minimum unless tested further using second order conditions.

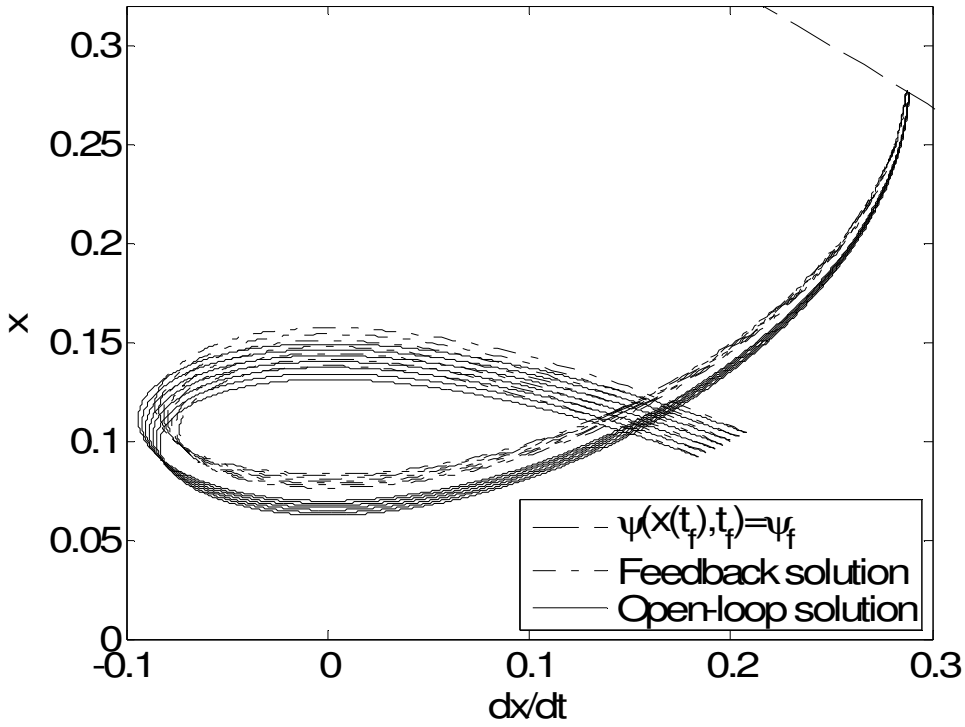


Fig.3.2 Case A: Phase Portraits with Various I.C. (Nominal I.C. Is Perturbed within the Range [-10 10] %)

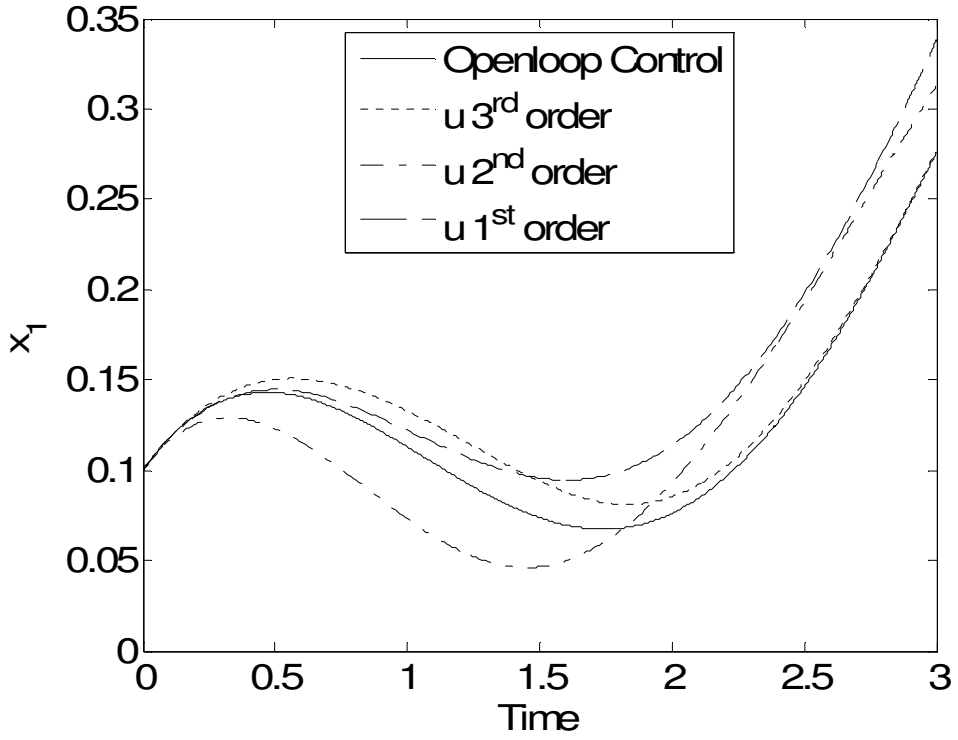


Fig. 3.3 Case A: Convergence of the State Trajectory as a Function of the Order of Approximation

Figure 3.4 shows three trajectories from the same initial conditions for the states, each with a different cost function. However, each feedback-controlled trajectory, for all the initial conditions, is associated with the lowest cost solution, for the respective case.

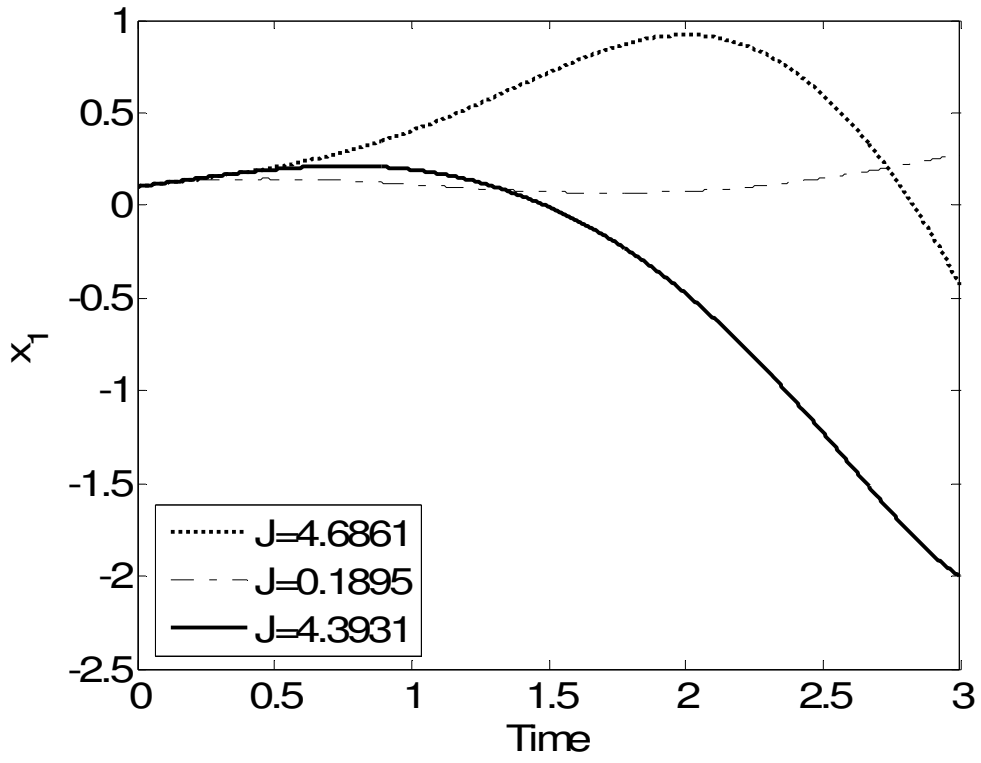


Fig. 3.4 Multiple Open-loop Extremals for the Same State I.C. but with Different Initial Costate Guesses

Case B: Unstable Nonlinear Plant and Nonlinear Terminal Constraint

Parameters for this example are given below.

$$
\begin{aligned}
&m_1 = -1; m_2 = 0.8; m_3 = 0.75 \\
&\varepsilon_1 = 0.9; \varepsilon_2 = 0.6; \varepsilon_3 = 0.6; \varepsilon_4 = 0.3; \varepsilon_5 = 0.1; \\
&Q_{11} = Q_{22} = B = R = 1; \\
&t_0 = 0; t_f = 3; x(t_o) = [0.1; 0.2] \\
&\psi_f = 0.5
\end{aligned}
\tag{3.5}
$$

The linear part of this system is unstable for this example. Figure 3.5 shows the phase portraits of the optimal trajectories, both closed-loop and open-loop. As can be seen, all the feedback and some of the open-loop trajectories, for neighboring initial conditions, lie close to each other. Some of the open-loop trajectories converge to higher-cost solutions, depending on the initial guesses for the costate values. As mentioned before, the same initial guess was used to obtain the solutions for all the initial conditions. To present the numerical elaboration on the achieved cost-to-go, the final states and the terminal accuracy obtained by using third order feedback solution and open-loop solution, respectively are demonstrated in Tables 3.1 and 3.2. For the perturbed initial conditions in the feedback solution, the same gains are utilized to achieve the minimizing terminal states.

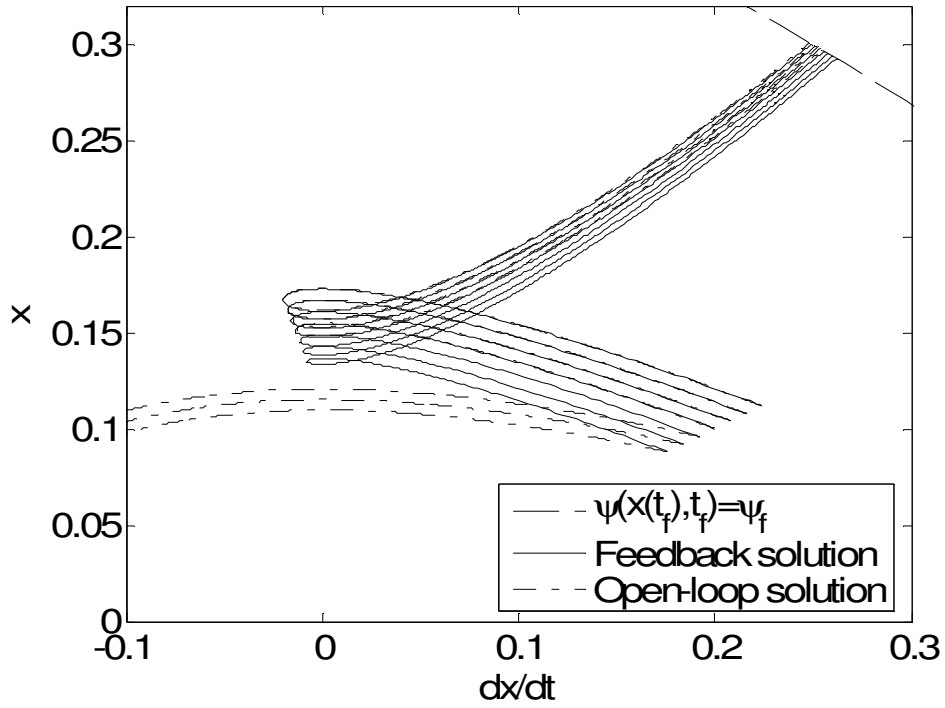


Fig. 3.5 Case B: Phase Portraits of the Feedback and Open-loop Solutions for Various State Initial Conditions

Table 3.1 CASE B: Higher Order Feedback Solution for 2-D Example

Optimal Feedback Third Order Series Solution						
Perturbation	x_1	x_2	***psi***	***%error***	***cost-J***	***final time***
-12%	0.30050	0.24886	0.50000378288000	-0.000756576	0.15326	2.9999996
-8%	0.29923	0.25095	0.50000210334000	-0.000420668	0.14505	2.9999911
-4%	0.29776	0.25302	0.49998881700000	0.002236599	0.13712	2.9999918
0%	0.29671	0.25508	0.49999731506000	0.000536988	0.12948	2.9999999
4%	0.29547	0.25771	0.49999900734000	0.000198532	0.12212	2.9999998
8%	0.29423	0.25913	0.49999666480000	0.000667040	0.1150	2.9999942
12%	0.29301	0.26113	0.50000794926000	-0.001589852	0.10826	2.9999987

Table 3.2 CASE B: Open-loop Solution for 2-D Example for Various I.C.

Open-loop Solution						
Perturbation	x_1	x_2	***psi***	***%error***	***cost -J***	***final time***
-12%	0.30085852	0.2483	0.500028365	-0.00567304	0.153250182	3
-8%	0.2995711	0.2504	0.500011	-0.00220005	0.145045045	3
-4%	0.29829133	0.2525	0.500007559	-0.00151178	0.137117379	3
0%	0.29701905	0.2546	0.500017793	-0.00355859	0.129470952	3
4%	-1.66789489	-1.9979	0.499988656	0.002268771	1.820155327	3
8%	-1.66921782	-1.9957	0.500028185	-0.005637	1.808490185	3
12%	-1.67053739	-1.9934	0.499991442	0.001711516	1.797130401	3

3.2 Application to Optimal Detumbling of Spacecraft

The SSM is applied to the problem of optimal detumbling maneuvers of a rigid asymmetric spacecraft. The Euler Equations for a rigid body are

$$\dot{\omega} = I^{-1}[\tilde{\omega} I \omega + u] \tag{3.6}$$

where $\omega = (\omega_1, \omega_2, \omega_3) \in \Re^{3x1}$ and $u=(u_1, u_2, u_3) \in \Re^{3x1}$, are the angular velocity and input control torque vectors, respectively. The moment of inertia matrix $I \in \Re^{3x3}$, is represented in the principal axes system as

$$I = \begin{bmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{bmatrix} \tag{3.7}$$

The angular velocity cross-product operator is represented by $\tilde{\omega}$. The principal moments of inertia for this example are given below[75]:

$$I_1 = 86.24 \text{ kg-m}^2; I_2 = 85.07 \text{ kg-m}^2; I_3 = 113.59 \text{ kg-m}^2 \tag{3.8}$$

The performance index is selected to be of the form

$$J = \frac{1}{2}\int_0^T (\omega^T Q \omega + u_1^{\,2} + u_2^{\,2} + u_3^{\,2}) dt \tag{3.9}$$

where Q is a positive definite weight matrix.

The solution presented is developed based upon an expansion of the cost-to-go to fourth order and a second order solution for the series reversion[66]. Series reversion to higher order was not carried out due to the excessive computational burden. A series reversion to second order was deemed sufficient for this problem after inspecting the results. All of the time-dependent gains were stored at discrete points of time during backward integration and the stored gains were utilized to calculate the feedback control and the trajectory, during the forward integration process. A cubic spline interpolation technique available in Matlab® [74] was used to determine gains at intermediate time points between two stored data points. Results obtained for a specific numerical example are presented below.

In the example considered, the final time was selected to be 2 sec and the weight matrix in the performance index is selected as the moment of inertia matrix, i.e., $Q = I$. The initial angular velocities selected are as shown below:

$$\omega_1(0) = -0.4 \text{ rad/sec}; \ \omega_2(0) = 0.8 \text{ rad/sec}; \ \omega_3(0) = 2 \text{ rad/sec}. \tag{3.10}$$

This example is not very realistic because of the high torque requirements but is used to demonstrate the accuracy of our methodology. Figure 3.6 shows the trajectories obtained by using the feedback solution and for comparison, the open-loop solutions are

also presented on the same graph. Again, there is a very close match between the two solutions. Figure 3.7 shows the control torques and slight deviations between the respective closed-loop and open-loop profiles are noticed. These deviations can be attributed to the very high initial angular velocities and the short maneuver time. Even though the differences between the respective control histories are noticeable, the open-loop and feedback controlled state trajectories are almost identical. As an additional check, the Lagrange multiplier ν, computed along the trajectory, is plotted in Fig. 3.8. Ideally, this vector should be constant but it is not so, due to series truncation. As mentioned previously, very near the final time, the Lagrange multipliers exhibit sharp changes. Hence the forward integration process was stopped just before reaching the final time.

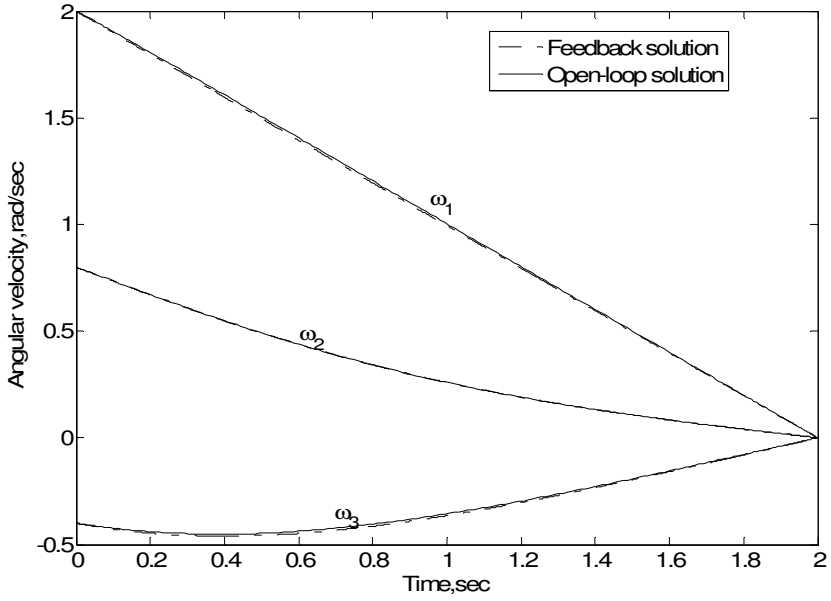


Fig. 3.6 Angular Velocities for the Spin Maneuver

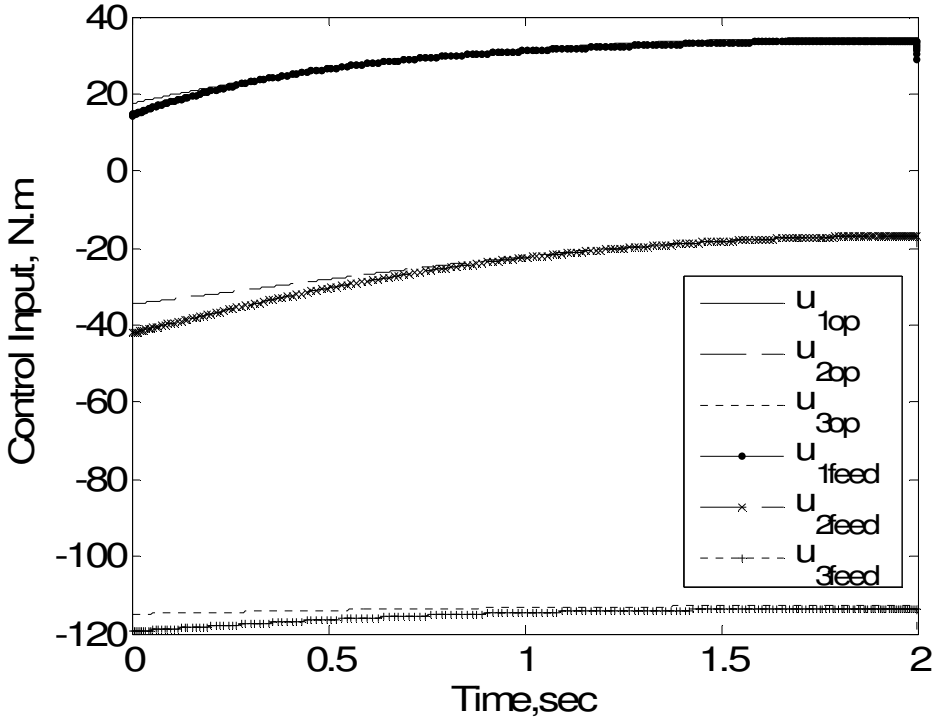


Fig. 3.7 Feedback and Open-loop Control Inputs

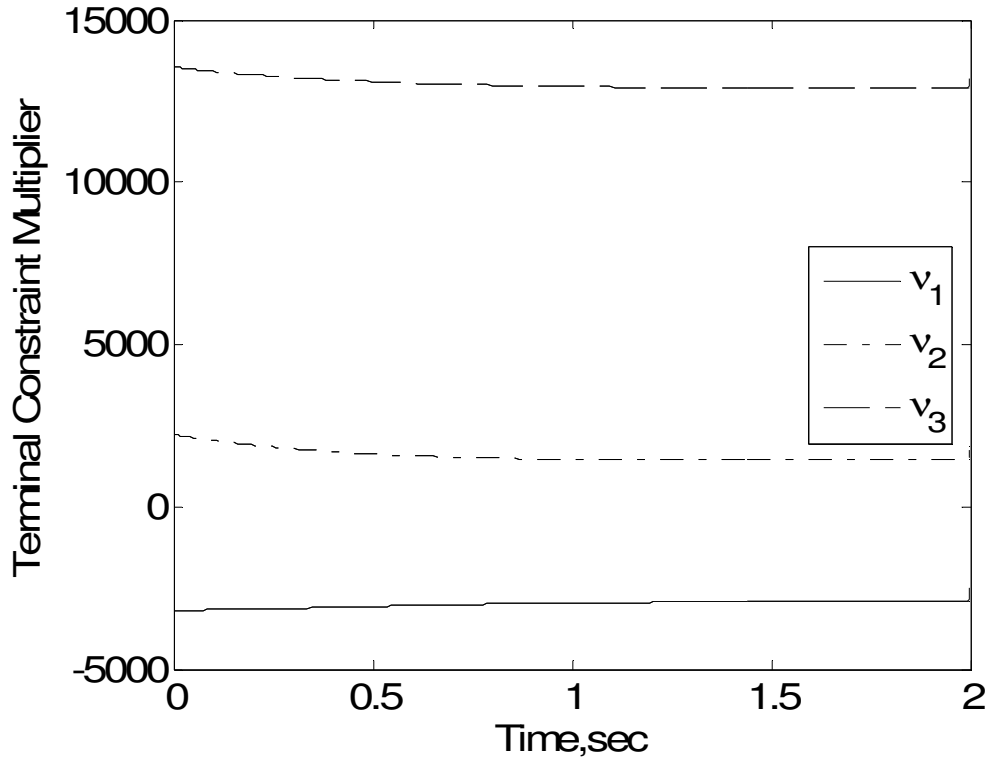


Fig. 3.8 Lagrange Multipliers Associated with the Terminal Constraints

3.3 Orbit Transfer Problem

The main problem considered in this chapter is the simultaneous design of an optimal trajectory and a guidance law for a minimum-fuel co-planar transfer of a spacecraft from an initial circular orbit to a specified final circular orbit, in a fixed time interval. It is assumed that the spacecraft is equipped with two thrusters which will operate continuously during the transfer. The performance index and the polar heliocentric equations of motion are given below:

Minimize:

$$J = \frac{1}{2}\int_{t_0}^{t_f} (u_2^2 + u_3^2)dt \tag{3.11}$$

subject to

$$\dot{r} = w; \tag{3.12}$$

$$\dot{v} = -\frac{wv}{r} + u_2 \tag{3.13}$$

$$\dot{w} = \frac{v^2}{r} - \frac{\mu}{r^2} + u_3 \tag{3.14}$$

where r is a radial distance from the sun, w and v are radial and tangential component of the velocity, respectively. The control inputs are specified as u_2 and u_3; μ is the gravitational constant. The angle θ, measured from a reference line is a free variable, governed by the following differential equation:

$$\dot{\theta} = \frac{v}{r}; \theta(t_0) = 0 \tag{3.15}$$

The given initial conditions are specified as shown below:

$$r(t_0) = R \tag{3.16}$$

$$v(t_0) = \sqrt{\frac{\mu}{R}} \tag{3.17}$$

$$w(t_0) = 0 \tag{3.18}$$

Boundary conditions at the final time specified functionally as $\psi(x(t_f)) = \psi_f$, are following;

$$r(t_f) = R_f \tag{3.19}$$

$$v(t_f) = \sqrt{\frac{\mu}{R_f}} \tag{3.20}$$

$$w(t_f) = 0 \tag{3.21}$$

Solution Methodology

The Hamiltonian for the system can be defined as

$$H = \frac{1}{2}(u_2^2 + u_3^2) + \lambda_r(w) + \lambda_v\left(\frac{-wv}{r} + u_2\right) + \lambda_w\left(\frac{v^2}{r} - \frac{\mu}{r^2} + u_3\right) \tag{3.22}$$

The states and costates are defined as $x = [r; v; w]$ and $\lambda = [\lambda_r; \lambda_v; \lambda_w]$, respectively. Minimization of the Hamiltonian w.r.t. the controls results in the following equations:

$$u_2 = -\lambda_v; u_3 = -\lambda_w \tag{3.23}$$

In order to apply the SSM, the Hamiltonian and the cost-to-go must be expanded about a reference optimal solution. The reference solution chosen is the trivial initial orbit, which requires no control to be maintained. Thus the reference values of the costates are all zeros and the reference states are constants. Thus the Hamiltonian can be written as,

$$H = \frac{1}{2}\begin{bmatrix}\delta x \\ \delta\lambda\end{bmatrix}^T \begin{bmatrix}H_{xx} & H_{x\lambda} \\ H_{\lambda x} & H_{\lambda\lambda}\end{bmatrix}\begin{bmatrix}\delta x \\ \delta\lambda\end{bmatrix} + ...H.O.T. \tag{3.24}$$

where

$$\delta x = \begin{bmatrix} r - R \\ v - \sqrt{\frac{\mu}{R}} \\ w \end{bmatrix} \tag{3.25}$$

and

$$\delta\lambda = \begin{bmatrix} \lambda_r \\ \lambda_v \\ \lambda_w \end{bmatrix} \tag{3.26}$$

The cost-to-go function can be expanded as

$$J = 0 + \frac{\partial J^{*}}{\partial x(t)}^{T} \delta x(t) + \frac{\partial J^{*}}{\partial \nu}^{T} \delta\nu + \frac{1}{2}\begin{bmatrix} \delta x \\ \delta\nu \end{bmatrix}^{T} \begin{bmatrix} J_{xx} & J_{x\nu} \\ J_{\nu x} & J_{\nu\nu} \end{bmatrix} \begin{bmatrix} \delta x \\ \delta\nu \end{bmatrix} + ...H.O.T. \tag{3.27}$$

The linear terms in the above equation vanish by consideration of Eqs. (2.11) and (2.15), and the reference solution adopted. Thus, the perturbed costates can be determined as follows:

$$\delta\lambda = \frac{\partial J}{\partial \delta x} = J_{xx}\delta x + J_{x\nu}\delta\nu + ...H.O.T. \tag{3.28}$$

Furthermore, the equivalent of Eq. (12) can be written as

$$\frac{\partial J}{\partial \delta\nu} = J_{\nu x}\delta x + J_{\nu\nu}\delta\nu + ...H.O.T. = 0 \tag{3.29}$$

Equation(3.28) can be substituted in Eq.(3.24) for the expansion of the Hamiltonian. The expanded Hamiltonian, together with Eq.(3.27), can be substituted in the HJB equation to proceed to the solution. A fourth-order power series for the cost-to-go function results in a third-order feedback control law. In this work, symbolic generation of the gain differential equations was achieved by using Maple and the gain differential equations were integrated by using the Matlab® routine RK45 and the solutions were stored at 201 equally-spaced time points. Equation(3.29) was solved for

δv using vector series reversion carried out to second order. The forward simulations of the closed-loop system were carried out by using a fixed-step fourth-order Runge-Kutta routine, RK4, available in Matlab®. Details of the numerical results are discussed in the next section.

Numerical Results

The initial and final conditions for the reference problem are given in Table 3.3 using non-dimensional units. The non-dimensional value of μ is 1.

Table 3.3 Initial and Final Conditions

	Initial time (t_0 =0)	Final time (t_f =2.4771 TU)
r (AU)	1	1.05
v(AU/TU)	1	0.9759
w(AU/TU)	0	0

where, $1\text{AU}=1.4959965\times10^{11}$ m ; $1\text{TU} = 58.132821$ days .

Figure 3.9 shows the polar plot of the optimal feedback trajectory from a circular orbit to another higher altitude orbit. Also plotted in this figure is the exact open-loop trajectory. One cannot distinguish between the feedback and open-loop solutions, given the scale of this figure. The error in achieving the desired radius is of the order of 1e-007 AU which corresponds to approximately 100 km. There are various reasons and implementation issues which contribute to the error. Foremost, there is the well-known "blind time"[3], during which there is a singularity in the computation of δv. In our

simulations, $\delta\nu$ was held constant near the final time. Furthermore, the Hamiltonian and the cost-to-go are expanded to fourth-order only and the forward integration used only 100 steps to reach the final time of 2.4771 TU. Another source of error is the second-order series reversion for the computation of $\delta\nu$.

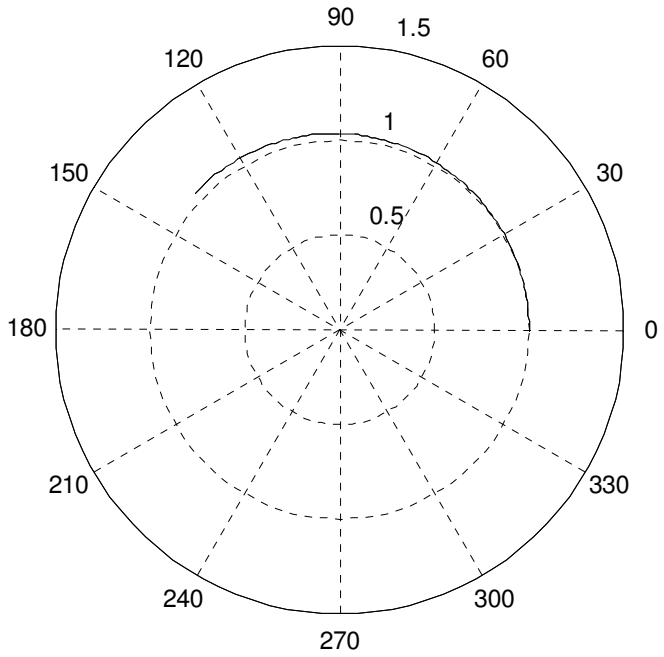


Fig.3.9 Orbit Transfer Using Higher Order Optimal Feedback Control

Fig. 3.10 and 3.11 show the errors between open-loop and feedback state trajectories with linear and nonlinear feedback input, respectively. The maximum error in the state variables while using the third-order control is limited to the range of $\pm 8e-6$. It is clearly demonstrated that the error levels diminish as the order of the control law is increased from linear to cubic. Figure 3.12 shows the control histories for different

orders of the control law. It is noticeable that feedback control of second-order deviates appreciably from the respective solutions for linear and third order feedback. Hence the truncation order should be chosen carefully. Figure 3.13 compares the open-loop and third order feedback control histories and there is very little difference between the two solutions.

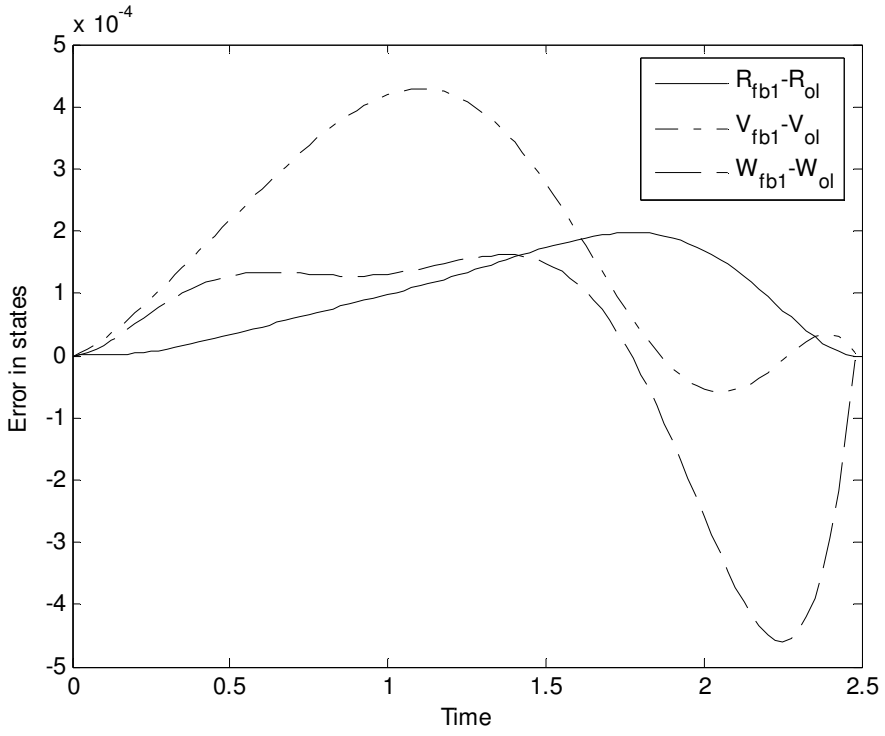


Fig. 3.10 Errors in the State Trajectories (Linear Feedback)

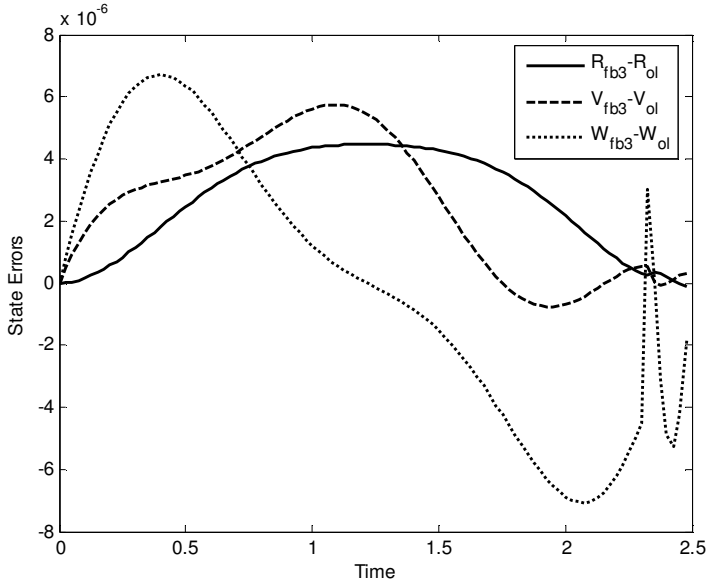


Fig. 3.11 Errors in the State Variable with Third-order Feedback Control

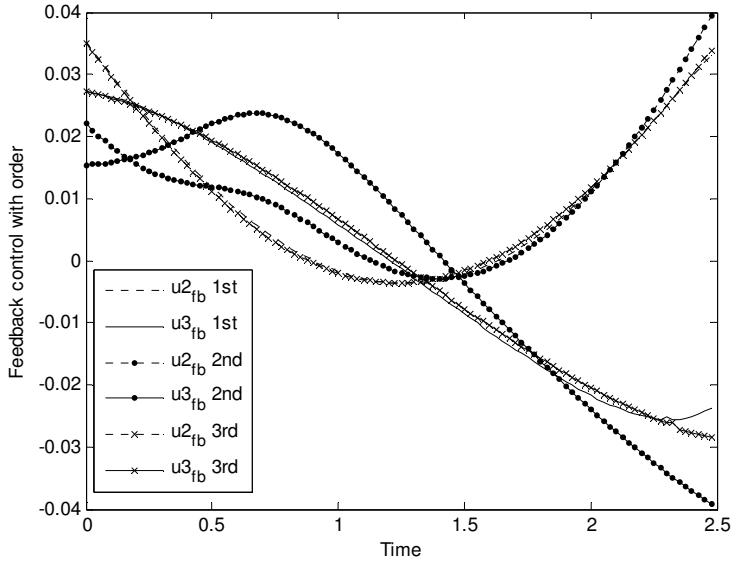


Fig. 3.12 Optimal Feedback Inputs for Various Orders of Feedback

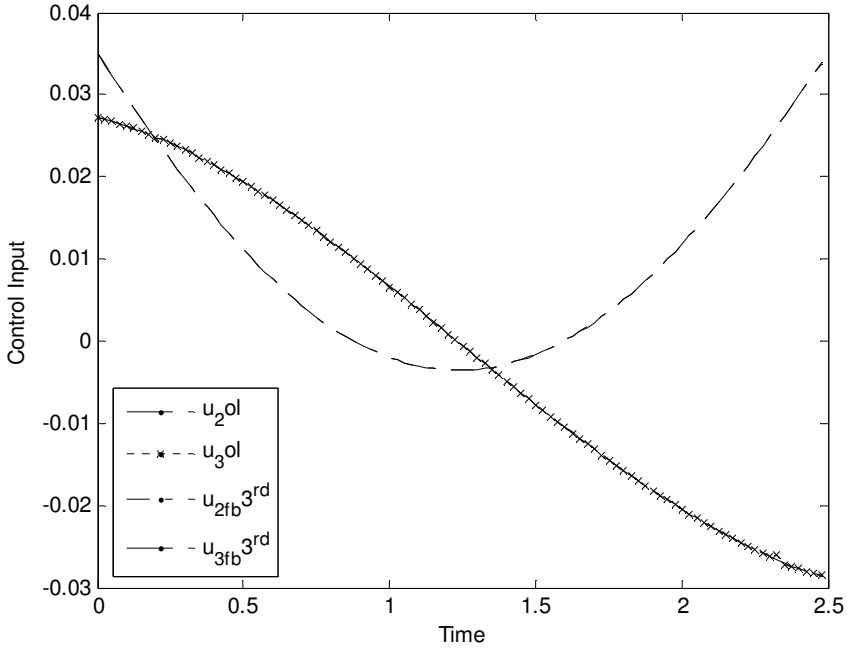


Fig. 3.13 Comparison of the Open-loop and Third-order Optimal Feedback Control Histories

The next set of figures show the robustness of the designed feedback control law and its ability to guide the spacecraft from perturbed initial conditions. Its sensitivity can be tested by perturbing either the initial condition or the final constraints. Here, the perturbation in the initial radius is considered in the range [0.998 1.01]. Figures 3.14-3.16 show the plot of the perturbed state trajectories converging to the terminal constraint. The terminal error is bounded within $\pm$ 5e-007 %, negligible in comparison to the errors in the initial conditions. Figure 3.17 shows the magnitudes of the applied feedback control inputs along the perturbed trajectories.

The variations of the terminal Lagrange multipliers are shown in Fig. 3.18. Even though these functions should theoretically remain constant throughout the trajectory, they are not so due to the truncation in the power series; their sharp variations near the end is due to the presence of the singularity discussed earlier. The cost-to-go functions for the perturbed trajectories are shown in Fig. 3.19. The bold line shows the actual cost with the nominal initial conditions. This plot shows that the cost-to-go and its time derivative along any particular trajectory are high near the initial and final times. There is a region during the mid-course flight where the cost-to-go is very nearly constant, indicating a coast region.

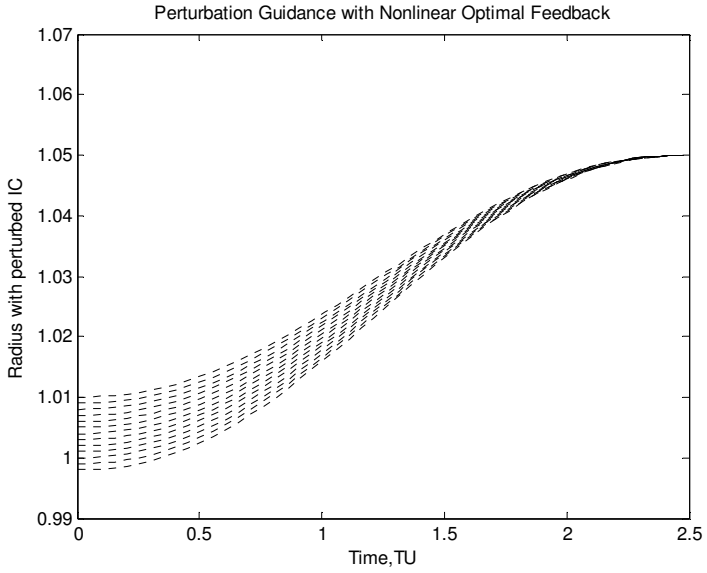


Fig. 3.14 Various Trajectories with Perturbed IC in 1% Range

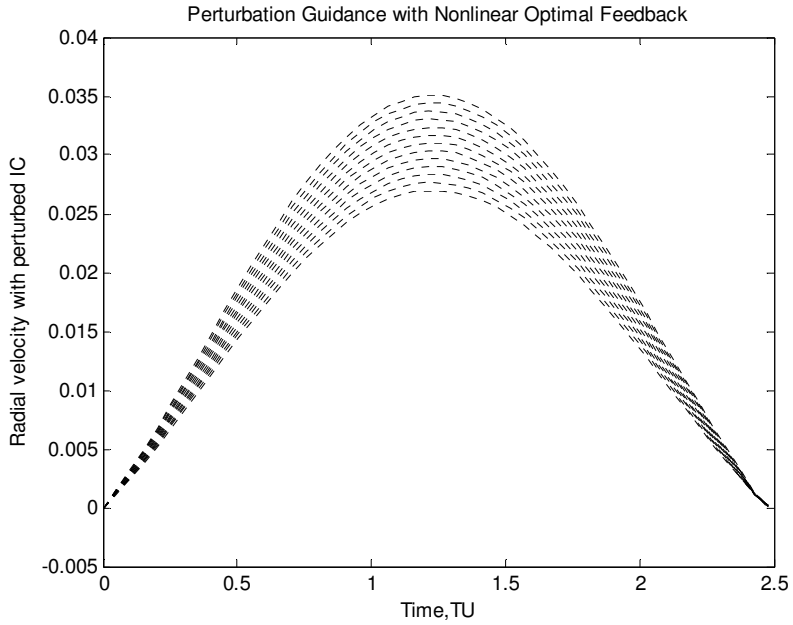


Fig. 3.15 Various Cases of Radial Velocity with Perturbed Radius in [0.998 1.01]

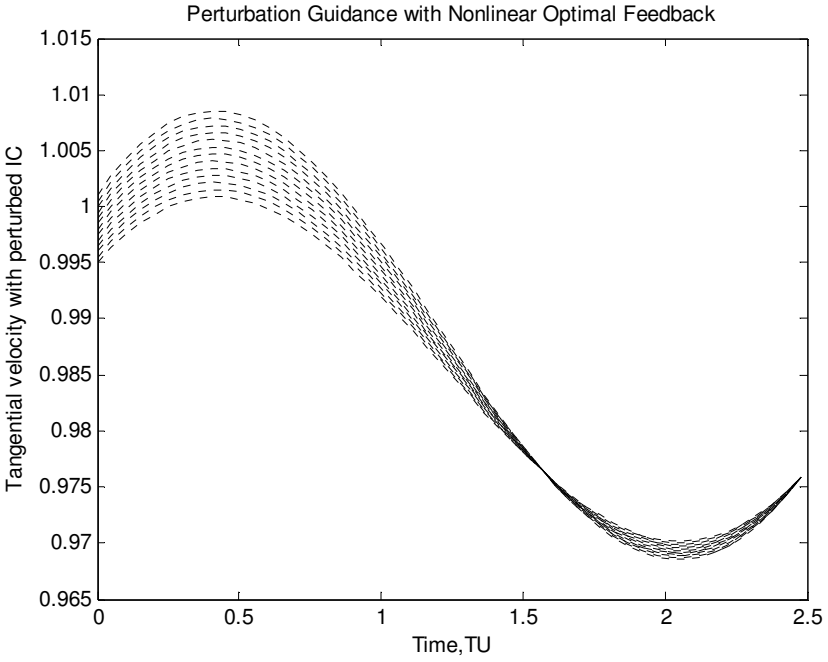


Fig. 3.16 Various Cases of Tangential Velocity with Perturbed I.C. in [-0.2 1] % Range

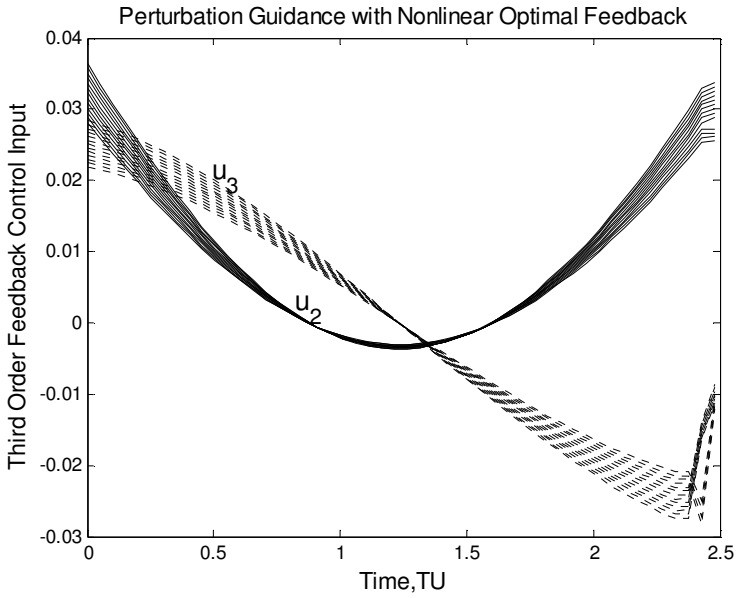


Fig. 3.17 Optimal Nonlinear Feedback for Perturbed Range of Initial Conditions

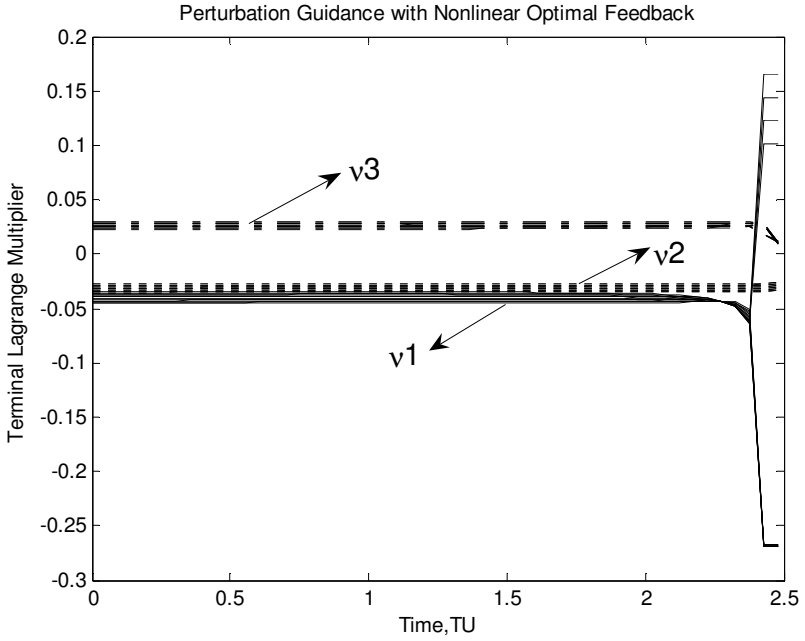


Fig. 3.18 Terminal Lagrange Multiplier for Perturbed Range of Initial Conditions

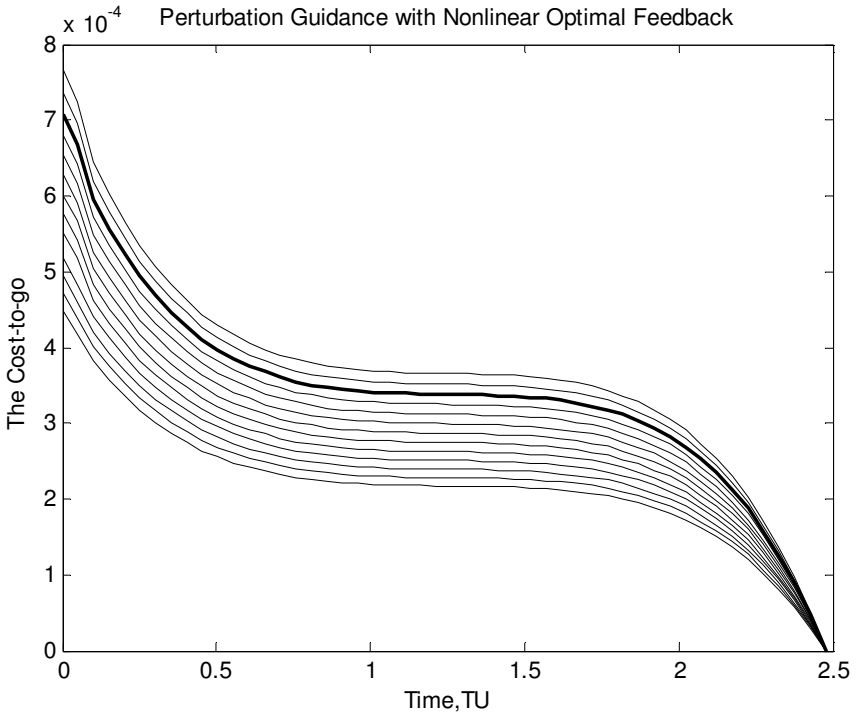


Fig. 3.19 Optimal Value Function for Perturbed Range of Initial Conditions

As a final example, the solutions for a 14-day transfer for the same nominal initial and final conditions are computed. Figures 3.20 and 3.21, respectively, show the errors in the state histories for linear and third-order feedback control. As can be seen, the errors decrease by two orders of magnitude with the introduction of the nonlinear feedback law. Figures 3.22 and 3.23, respectively, show the control errors for the linear and nonlinear guidance laws. A reduction of error by three orders of magnitude can be seen in the third-order control applied to the 14 days orbit transfer problem.

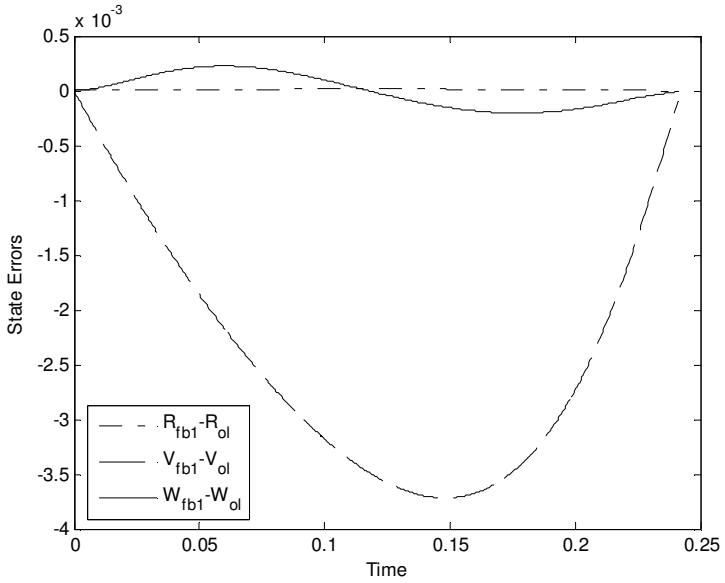


Fig. 3.20 State Errors in Optimal Open-loop and Feedback Trajectories (Linear Feedback)

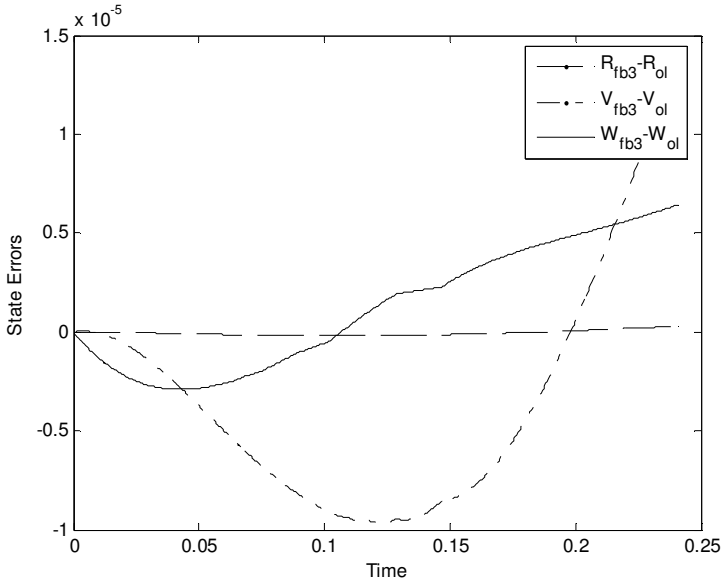


Fig. 3.21 State Errors in Optimal Open-loop and Feedback Trajectories (Third-order Feedback)

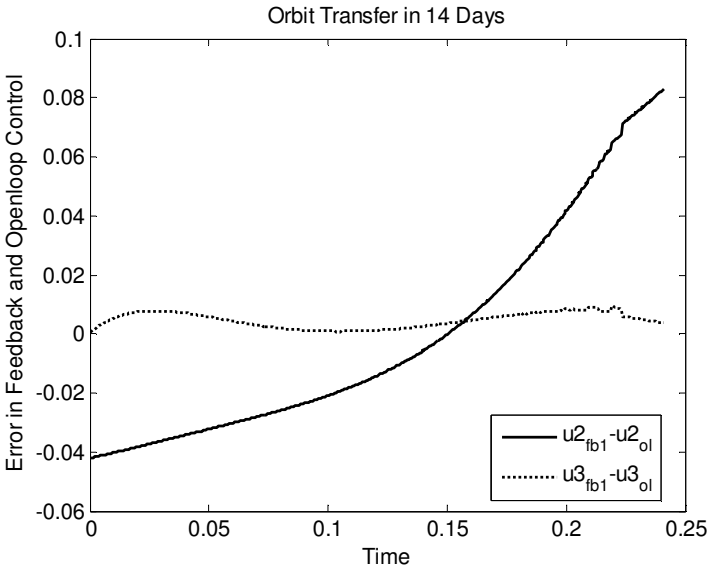


Fig. 3.22 Errors in Optimal Open-loop and Linear Feedback Control Histories

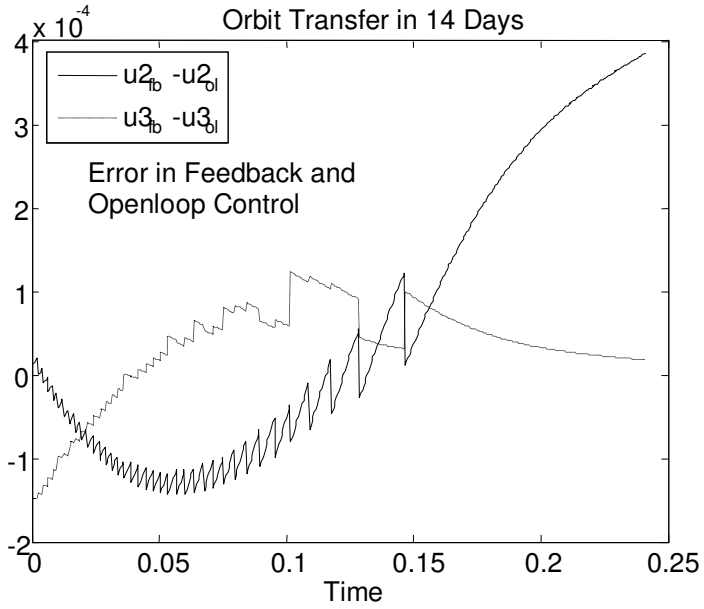


Fig. 3.23 Errors in Optimal Open-loop and Third-order Feedback Control Histories

CHAPTER IV

OPTIMAL NONLINEAR FEEDBACK CONTROLLER DESIGN USING A WAYPOINT SCHEME

This chapter presents a waypoint scheme, designed to broaden the scope of SSM. It has been shown previously that the inevitable truncation of the series and the process of vector series reversion required for calculating the constraint Lagrange multipliers can cause significant deviations in the trajectory from the optimal, especially during the midcourse. The validity of the Taylor series expansion of a given order also depends on the nonlinearity of the system and the choice of the final time. In order to avoid such divergence, the assumed polynomial expansion often requires a large number of terms. Also, in general, to obtain a high level of accuracy, the dynamic programming approach requires a voluminous storage of gains, especially, if the problem is multidimensional and solved over a large time domain. Therefore, the waypoint scheme, involving partitions of the time domain, is suggested to overcome these limitations and obtain the optimal feedback solution with a reduced computational burden.

The efficacy of the method is demonstrated on two example problems:

1. A highly nonlinear two dimensional polynomial dynamic system.
2. A low-thrust orbit transfer problem.

4.1 The Waypoint Scheme

Introduction of waypoints converts a large time-domain dynamic optimization problem into many interconnected, smaller time-domain optimal control sub problems. The waypoint scheme imposes constraints to establish connections to preserve optimality of the solution with respect to the original problem. The idea behind the waypoint scheme is to implement the SSM in small time segments, thereby diminishing the errors in the midcourse trajectory and enhancing the solvability of the HJB equation.

The waypoint scheme presented in this dissertation divides the given time domain into smaller finite-time intervals of equal length. For autonomous systems, the gains need only be calculated over a single interval. The gain obtained over one interval can be reused in the other intervals. With the known gains, a parameter optimization problem is constructed at segment boundaries to obtain the optimal waypoint state variable specifications. A waypoint specification acts as a terminal constraint vector for the segment to its immediate left. The waypoint algorithm is especially attractive for problems which cannot be handled by using a finite number of terms in the power series expansion or for which the series reversion process diverges during the forward integration. To provide a clear exposition of the waypoint scheme, we consider a theoretical example with a single waypoint. Without loss of generality, the scheme can be applied for problems with several waypoints.

A single waypoint at time t_1 is introduced by partitioning the time domain, $[t_0, t_f]$ into two equal parts: $[t_0, t_1]$ and $[t_1, t_f]$. Thus, two terminally constrained OCPs

equivalent to the one defined as Eq. (2.1) are obtained. However, the equivalence mentioned above is true if the state variable constraint at the waypoint is specified such that it lies on the optimal trajectory resulting from the original problem. In general, this constraint value has to be determined to satisfy the criteria for optimality. The intermediate constraint at the waypoint is specified as follows:

$$\psi(t_1) \equiv x(t_1) - x_1 = 0 \tag{4.1}$$

where, x_1 is an unknown to be determined. As such, choosing the waypoint location at the midpoint is not a requirement, but it allows for a significant reduction in the storage of gains, especially for autonomous dynamical systems.

4.2 The Waypoint Computation

In this section, a necessary condition for determining the optimal waypoint constraint value is presented. If this constraint is chosen arbitrarily, then the total cost of following the trajectory from the initial point to the terminal point, via the selected waypoint, will be higher than the optimal cost. The least cost is obtained if the chosen waypoint lies on the optimal trajectory of the original problem.

SSM assumes a higher-order polynomial series representation of cost-to-go, J^* in terms of $x(t)$ and terminal Lagrange multipliers ν with an unknown gain. The expansion for J^*, using indicial notation is,

$$J^*(x(t),t) = S_{1_{ij}}(t)x_i x_j + S_{2_{ijk}}(t)x_i x_j x_k + .. + P_{1_{pj}}(t)\nu_p x_j + P_{2_{pqj}}(t)\nu_p \nu_q x_j +$$
$$P_{3_{pij}}(t)\nu_p x_i x_j + ... - (\psi_f)_p \nu_p + V_{1_{pq}}(t)\nu_p \nu_q + V_{2_{pqr}}(t)\nu_p \nu_q \nu_r + ...H.O.T \quad (4.2)$$
$$i, j, k, l.....etc. = 1, 2, 3, ..., n, \quad p, q, r.....etc = 1, 2, 3, ..., p \le n$$

where $S_i(t), P_i(t), V_i(t), ...,$ i=1,2,3,.. , are time-dependent gain tensors.

As derived in chapter II, the key equations, which connect J^* with λ and ν, are,

$$\lambda(t) = \frac{\partial J^*(x(t),t)}{\partial x(t)} \quad (4.3)$$

$$\frac{\partial J^*(x(t),t)}{\partial \nu} = 0 = (\psi(x(t_f)) - \psi_f) \quad (4.4)$$

An advantage of SSM is that the state boundary conditions do not affect the gain differential equations. The optimal cost-to-go at an instant *t* can be expressed as a function of the instantaneous state, terminal Lagrange multiplier, and time dependent gains as,

$$J^* = J^*(x(t), x_1, \nu(x, x_1), gains(t)) \quad (4.5)$$

Since the optimal waypoint constraint specification is required to minimize the total cost of the trajectory, it must satisfy the necessary condition

$$\frac{dJ^*}{dx_1} = 0 \quad (4.6)$$

The above necessary condition results in a parameter optimization problem which can be solved, in the outer loop, analytically. In the following section, other necessary conditions which must be applied together with the condition given above are presented for the case of one waypoint.

4.3 Derivation of Necessary Conditions

Figure 4.1 shows a sketch of the optimal trajectory with respect to time as well as feasible trajectories obtained for various choices of the waypoint constraint value. Let the waypoint be located at *A* and let *O* and *B* denote the given initial and final conditions at t_0 and t_f, respectively.

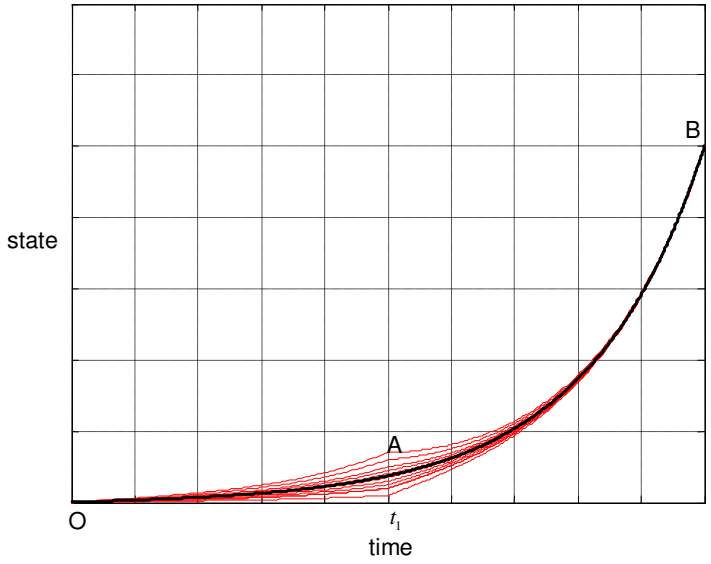


Fig. 4.1 Illustration of One Waypoint along the Trajectory

The objective is to minimize the total cost, J^*_{OB} with respect to $x_1(t_1)$. The total cost J^*_{OB} can be written as,

$$J^*_{OB} = J^*_{OA} + J^*_{AB} \tag{4.7}$$

and the application of the necessary condition of Eq.(4.6) results in

$$\frac{dJ^*_{OB}}{dx_1} = \frac{dJ^*_{OA}}{dx_1} + \frac{dJ^*_{AB}}{dx_1} = 0 \tag{4.8}$$

$$\Rightarrow \frac{\partial J^*_{OA}}{\partial x_1} + \frac{\partial J^*_{OA}}{\partial \nu_O}\frac{\partial \nu_O}{\partial x_1} + \frac{\partial J^*_{AB}}{\partial x_1} + \frac{\partial J^*_{AB}}{\partial \nu_A}\frac{\partial \nu_A}{\partial x_1} = 0 \tag{4.9}$$

where $v_O = \nu(t_1^-)$ and $\nu_A = \nu(t_1^+)$, are the values of the terminal Lagrange multipliers in segments OA and AB, respectively.

Substituting appropriate forms of Eq. (4.4) in the above equation, i.e.,

$$\frac{\partial J^*_{OA}}{\partial \nu_O} = \frac{\partial J^*_{AB}}{\partial \nu_A} = 0 \tag{4.10}$$

results in the following:

$$\Rightarrow \frac{\partial J^*_{OA}}{\partial x_1} + \frac{\partial J^*_{AB}}{\partial x(t_1^+)} = 0 \tag{4.11}$$

It is assumed that J^* is a smooth and continuous function of x.

From Eq.(4.2), it is known that J^*_{OA} is a linear function of the waypoint constraint value, x_1 which implies that,

$$\frac{\partial J^*_{OA}}{\partial x_1} = -\nu_A(t_1^-) \tag{4.12}$$

Then, continuing from Eq. (4.11)

$$\Rightarrow -\nu_A(t_1^-) + \frac{\partial J^*_{AB}}{\partial x(t_1^+)} = 0 \tag{4.13}$$

Now, in segment AB, application of the transversality condition of Eq. (4.3) results in the following:

$$\lambda(t_1^-) = \lambda(t_1^+) \tag{4.14}$$

The above analysis proves that the optimal waypoint specification is determined to ensure the continuity of the costates as well as the Lagrange multipliers at the segment boundary.

4.4 The Complete Procedure with Multiple Waypoints

The waypoint method is a feedback equivalent of the multiple shooting method[67]. It is a two-loop optimization process characterized by inner and outer loops. The inner loop is responsible for the implementation of the feedback law, whereas the outer loop deals with the optimization of the waypoint location. Before proceeding to the numerical examples, this section summarizes the overall procedure to compute the feedback solution in the presence of multiple waypoint constraints:

1. Partition the given time-domain into smaller segments to locate the waypoints and choose the waypoint specification.
2. Apply the series solution methodology to obtain the gain differential equations.
3. Integrate backwards in the last time interval to store gains.
4. In the outer loop, apply the static optimization procedure mentioned above to compute the optimal values of the waypoint states.
5. Once the waypoints are known, the closed-loop system can be simulated in conjunction with the higher-order feedback control law. In the first segment,

the first waypoint acts like a terminal constraint; in the next subsequent segment, the previous waypoint is treated as an initial condition and the next waypoint becomes a terminal constraint.

In the subsequent sections, two numerical examples are considered to demonstrate the usefulness of the waypoint based SSM.

4.5 A 2-D Nonlinear Dynamical System

To obtain the feedback control law by minimizing the energy of a 2-D nonlinear system subject to the point terminal constraints**,** the optimal control problem (OCP) is given as,

Minimize:

$$J = \frac{1}{2}\int_0^3 (x^2 + \dot{x}^2 + u^2)dt \tag{4.15}$$

subject to

$$\begin{aligned} &\ddot{x} = -1.8x^2 + 0.75x^3 + u; \\ &x(0) = 0.1; \dot{x}(0) = 0.2; \\ &\psi \in \Re^{2\times 1} \equiv x(3) = 0.5,\ \dot{x}(3) = 0.5 \text{ (terminal constraints)} \end{aligned} \tag{4.16}$$

Solution Procedure

For the given problem, the HJB equation is,

$$-\frac{\partial J^*}{\partial t} = \underbrace{\min}_{u}\left[\frac{1}{2}(x^2+\dot{x}^2+u^2)+\frac{\partial J^*}{\partial x}(\dot{x})+\frac{\partial J^*}{\partial \dot{x}}(-1.8x^2+0.75x^3+u)\right], \qquad (4.17)$$
$$J^*(x(t_f),t_f)=0$$

By using the necessary conditions[23] for optimality, the optimal feedback control can be written as,

$$u^* = -\frac{\partial J^*}{\partial \dot{x}} \qquad (4.18)$$

Note that there is no linear term in x, in Eq.(4.16), which makes the OCP difficult to solve using SSM. Proceeding with the straightforward expansion of J^* to sixth-order, it is found that the series reversion process, which computes the value of ν during the forward integration, results in a significant error in the computed control. Thus, the usual single-segment procedure is deemed unsuitable to handle the strong nonlinearity present in this system.

Next, the above OCP is solved by using two waypoints, at $t_i = 1, 2$ sec. As before, J^* is assumed to be a polynomial series of sixth order in the states and second order in ν. It contains a total of 115 gains. The gains are computed for the last time interval: $t = [2, 3]$ .Additional necessary conditions are utilized (outer loop) to obtain the optimal waypoints at these time instants.

Numerical Results

The waypoint selection methodology in conjunction with the series solution results in the following:

$$x(1) = 0.14452004365, \quad \dot{x}(1) = -0.030242702365 \tag{4.19}$$

$$x(2) = 0.17207973378, \quad \dot{x}(2) = 0.1371940561405 \tag{4.20}$$

Upon forward integration with the closed-loop controller in place, the open-loop and feedback costs are obtained as shown below:

$$J_{OL}^{*} = 0.3266422, \quad J_{FB}^{*} = 0.3269268 \tag{4.21}$$

In order to demonstrate the outer loop optimization process, several waypoint values in the vicinity of the optimal are selected. Figures 4.2 shows the variation of J^{*} at t=1 and 2 seconds with respect to the arbitrary waypoint values, from which the existence and locations of the optimal waypoints can be ascertained. Figure 4.3 shows the trajectories of x_2 for the various waypoint values chosen. The optimal end-to-end trajectory is shown by the dashed line. Figure 4.4 presents several phase plots obtained by using arbitrary waypoint values at $t_i = 1, 2\,\text{sec}$. The bold line shows the open-loop solution, which is associated with the optimal waypoints. The computed feedback solutions with arbitrary waypoint values are shown in red. The cost for the trajectory with waypoint values located on the open-loop trajectory is the least among the other trajectories shown.

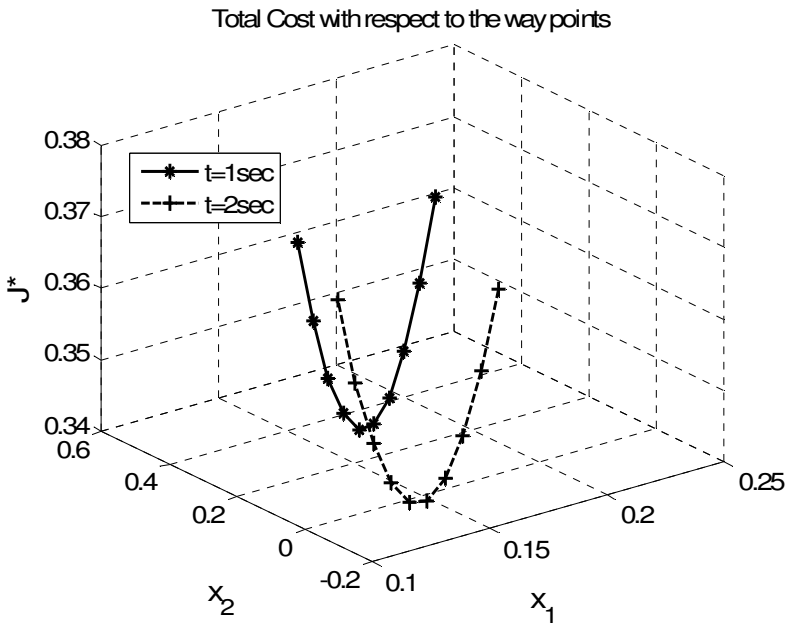


Fig. 4.2 Cost at t=1 and 2 with Arbitrary Waypoints

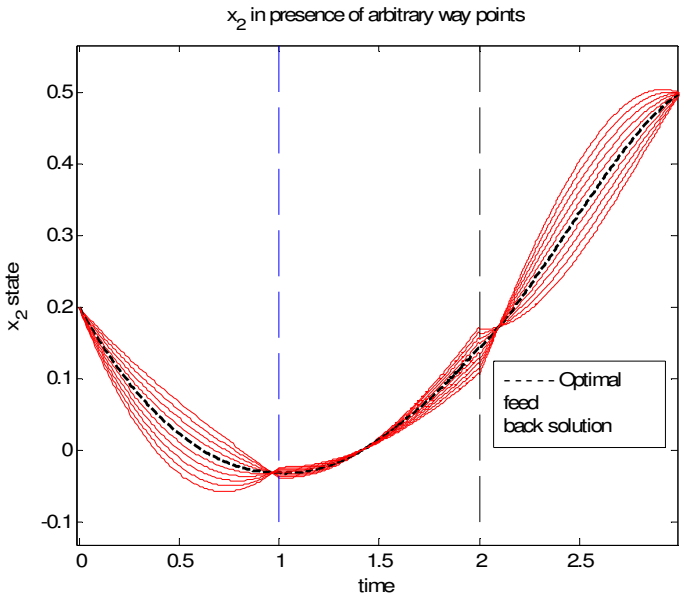


Fig. 4.3 x_2 at t=1 and 2 with Arbitrary Waypoints

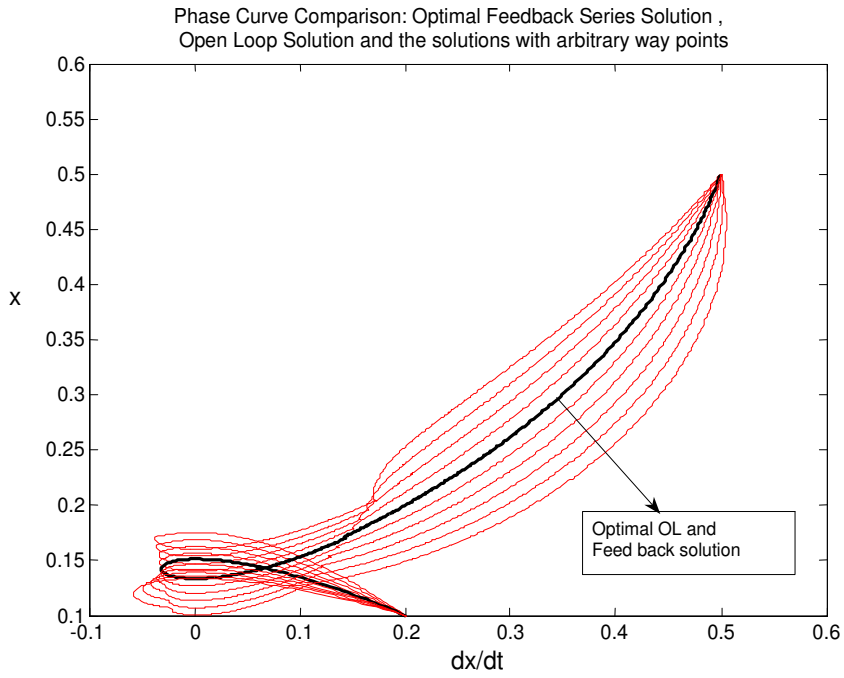


Fig. 4.4 Phase Curve Comparison with Arbitrary Waypoints at t=1 and 2

Figure 4.5 shows the final solution obtained by using the optimal waypoints. The open-loop and feedback solutions obtained using two waypoints are nearly indistinguishable. The computed solution obtained without using waypoints is also shown in this figure and it shows considerable error in the midcourse. Figure 4.6 shows the cost-to-go comparison for the open-loop and feedback controls with and without waypoints.

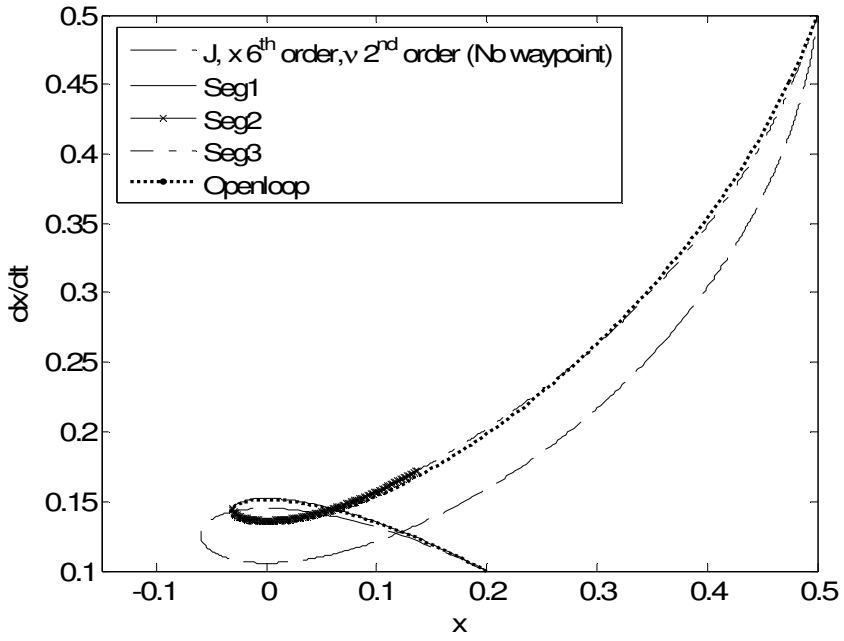


Fig. 4.5 Series Solution with 2 Waypoints

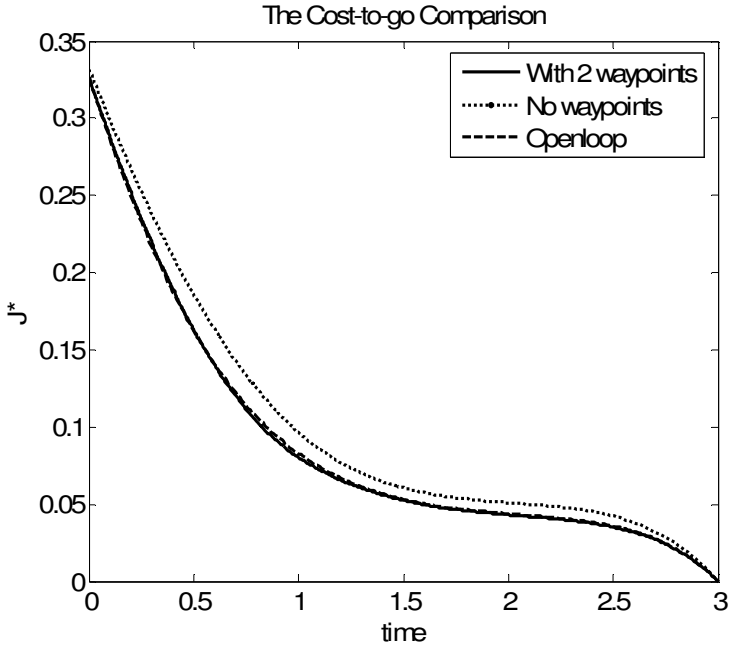


Fig. 4.6 The Cost-to-go Comparison with Open-loop Solution

4.6 Minimum Fuel Earth to Mars Orbit Transfer

The next example deals with a standard orbit transfer problem that has been considered by many researchers. The objective here is to validate the performance of the waypoint scheme on the problem of minimum-fuel orbit transfer from Earth to Mars in 144 days. A detailed analysis of the same example was presented in the previous chapter. For the sake of completion, the OCP can be stated again as,

Minimize:

$$J = \frac{1}{2}\int_0^{t_f} (u_2^2 + u_3^2)dt \tag{4.22}$$

subject to

$$\dot{r} = w; \dot{v}_T = -\frac{wv_T}{r} + u_2; \dot{w} = \frac{v_T^2}{r} - \frac{\mu}{r^2} + u_3 \tag{4.23}$$

$$r(0) = R = 1AU; v_T(0) = \sqrt{\frac{\mu}{R}}; w(0) = 0 \tag{4.24}$$

Terminal constraints: $\psi \in \Re^{3\times 1} \equiv r(t_f) = R_f = 1.54AU; v_T(t_f) = \sqrt{\frac{\mu}{R_f}}; w(t_f) = 0$ (4.25)

$t_f = 144$ days, 1AU=1.4959965×10^{11} m ;1TU = 58.132821 days

In the preceding equations, r, v_T and w, denote the radial position, the tangential velocity component and the radial velocity component, respectively. u_2 is the control along the tangential direction and u_3 is the control along the radial direction. The constant μ is the gravitational parameter of the sun.

Solution Procedure

The Hamiltonian *H* can be written as,

$$H = \frac{1}{2}(u_2^2 + u_3^2) + \lambda_r(r) + \lambda_{v_T}(-\frac{v_T w}{r} + u_2) + \lambda_w(\frac{v_T^2}{r} - \frac{\mu}{r^2} + u_3) \tag{4.26}$$

Since the series solution method requires a polynomial form of the plant dynamics, the Hamiltonian of the system can be expanded about a reference orbit. In this solution, a circular orbit of the Earth radius, *R* is considered as a reference. Hence, *H* can be specified in terms of the new co-ordinates defined below:

$$\delta x = \begin{bmatrix} r - R \\ v_T - \sqrt{\frac{\mu}{R}} \\ w - 0 \end{bmatrix} \tag{4.27}$$

The HJB equation for this example is given below:

$$\frac{\partial J^*}{\partial t} = -\underbrace{\min}_{u_2, u_3}\left[H_1(\delta x, u_2, u_3, \lambda_r, \lambda_{v_T}, \lambda_w) \right] \tag{4.28}$$

where H_1 is a fourth-order Taylor series expansion of *H* with respect to δx.

By using the necessary condition, $\frac{\partial H_1}{\partial u_{2,3}} = 0$; the corresponding optimal controls are obtained as

$$u_2 = -\frac{\partial J^*}{\partial \delta x_2}; u_3 = -\frac{\partial J^*}{\partial \delta x_3} \tag{4.29}$$

Numerical Results and Discussion

Since the OCP is posed with hard terminal constraints, J^* is expanded in a power series involving ν and δx. The regular series solution algorithm suffers from errors in the series reversion process due to the low order of the expansion about a circular orbit. Based upon the given transfer time, the computational difficulty of series reversion can be avoided by considering a higher order expansion of the cost-to-go for all the states but the expansion order for ν is limited to second order only. In such a case, the series reversion process is easy to implement but it results in a significant error in the midcourse trajectory, which in turn leads to a higher cost-to-go. To bypass such problems with the assumed expansion of J^*, the waypoint scheme is employed to obtain more accurate results. The HJB equation is solved by using a sixth-order series expansion in x and second-order in ν, with 2, 4 and 12 equally spaced waypoints and the obtained solutions are compared with the open-loop solution. The following table shows the values of J^* for the different cases.

Table 4.1: The Cost-to-go Analysis for the Orbit Transfer Problem

J^*_{OL} [By using shooting method]	J^*_{FB} [Sixth order in x and 2nd order in ν, no waypoint]	J^*_{FB} [Higher order feedback with 2 waypoints]	J^*_{FB} [Higher order feedback with 4 waypoints]	J^*_{FB} [Higher order feedback with 12 waypoints]
0.09182	0.131412	0.10512	0.098204	0.094162

Figure 4.7 demonstrates polar plots of the transfer trajectory obtained by using three different methods: open-loop, series solution without waypoints and series solution with two waypoints. Without the use of waypoints, the transfer follows a completely different midcourse trajectory from that of the open-loop solution. However, the inclusion of 2 waypoints provides a significant correction to the required control and midcourse trajectory as well. Figs. 4.8-4.9 show the transfer orbit with 4 and 12 waypoints. As the number of waypoints is increased from 4 to 12, the higher-order feedback solution depicted by the dotted bold line tends closer to the open-loop solution, shown by the dashed line. The cost-to-go values in Table 4.1 show that inclusion of more waypoints improves the midcourse solution. However, it should be noted that the choice of the total number of waypoints depends upon the OCP and the nonlinearity of the dynamical system. The deviation in the midcourse trajectory can be reduced up to a certain extent by using a judicious combination of the number of waypoints and the number of terms used in the series expansion.

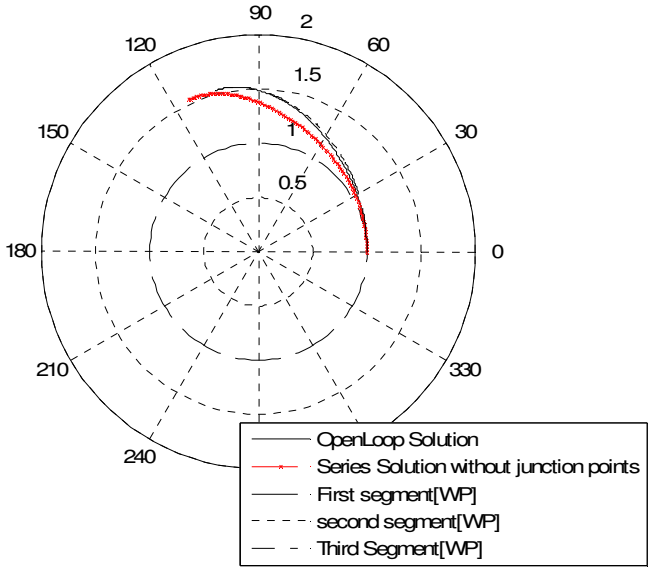


Fig. 4.7 A Comparative Study: Series Solution, Series Solution with 2 Waypoints and Open-loop Solution

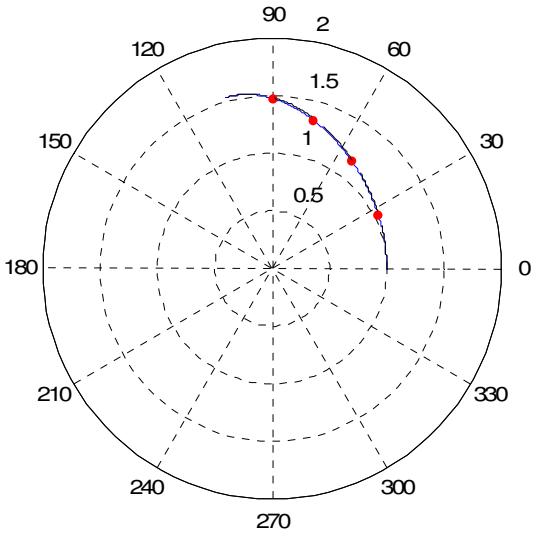


Fig. 4.8 Higher Order Feedback Solution with 4 Waypoints

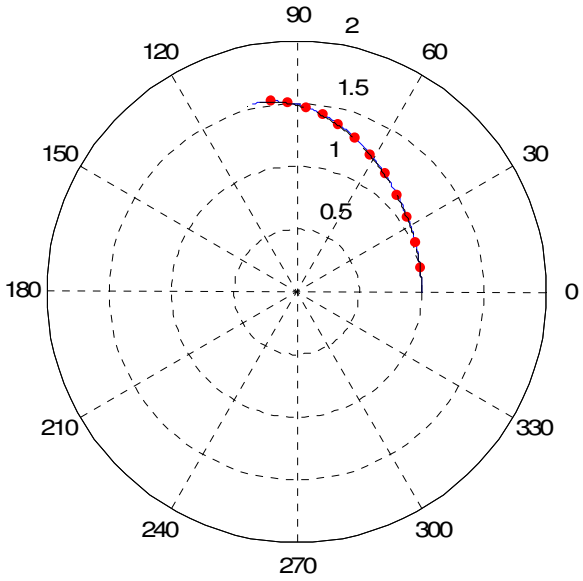


Fig. 4.9 Higher Order Feedback Solution with 12 Waypoints

It is well-known that ν becomes singular very close to the final time. Using many waypoints and short intervals exacerbates this singular behavior. It is found that considering 32 waypoints in the same orbit transfer problem increases the cost-to-go to 0.0966. Even if the computation of ν is stopped just before the terminal time, it causes an error in the waypoint calculation. Also, shrinking the time-domain for each segment by increasing the number of waypoints does not allow the same gains to be used in each interval.

CHAPTER V

A SERIES SOLUTION METHOD FOR THE SOLUTION OF THE HAMILTON JACOBI ISAACS EQUATION AND ITS APPLICATIONS TO AEROSPACE SYSTEMS

This chapter is primarily focused on developing the approximate analytical solution of the finite-time Hamilton Jacobi Isaacs (HJI) equation with soft and hard terminal constraints. Like the HJB equation, the HJI is a first order nonlinear partial differential equation. Analytical solutions to the HJI equation cannot be obtained in general. In fact, approximation of the HJI solution by using numerical methods is a challenging task because the admissible solution (unlike the HJB solution) exists in the game-theoretic saddle region

The first section of the chapter covers a brief introduction of the HJI equation associated with nonlinear dynamics, followed by the application of SSM to solve it. Subsequent sections present some examples which are mainly considered to show the application of SSM to solve the HJI equation. The first important application is in the area of dynamic games. A two-player nonlinear pursuit-evasion example involving satellites in orbits is considered. The optimal control problem is posed with a soft terminal constraint for the desired terminal capture. The study of this example is motivated by a need for space situational awareness. It may be required to conduct a

non-contact proximity sensing of one spacecraft moving in a different orbit by another spacecraft for health monitoring or identification purposes.

As has been mentioned earlier, another important application problem is H_2 / H_∞ design for terminally constrained problems. The subsequent section presents H_∞ feedback laws for a three dimensional short period pitch dynamics of a missile posed on a finite horizon with point terminal constraints. First, the open-loop solution is given for a rest-to-rest maneuver to show the effect of the attenuation factor, γ. Then, for a fixed value of γ, a third-order feedback solution is obtained by employing the SSM. The results obtained are compared with the respective open-loop solutions. For both the applications, several examples are illustrated to elucidate the methodology and its robustness to uncertain initial conditions.

5.1 Overview of the HJI Equation

In the area of differential dynamic games, due to the historical contribution by Isaacs[1], the equation for generating the field of extremals is known as the Hamilton Jacobi Isaacs (HJI) equation. This section, briefly overviews the formulation of the HJI equation with the Series Solution Methodology to solve a finite-time general optimal control problem in the feedback form.

The HJI Equation

Consider a general optimal control problem stated as

Minimize the following performance index with respect to u and maximize it with respect to w:

$$J = \phi(x(t_f)) + \int_{t_0}^{t_f} L(x,u,w,\tau)d\tau \tag{5.1}$$

subject to the nonlinear dynamic constraint,

$$\dot{x} = f(x,u,w,t); x(t_0) = x_0 \tag{5.2}$$

with a terminal constraint specified at a fixed final time t_f:

$$\psi(x(t_f)) - \psi_f = 0;\ \psi, \psi_f \in \Re^{p \le n} \tag{5.3}$$

where f: $\Re^{n+m+s+1} \to \Re^n$ is a smooth, analytic, vector-valued function with $x \in \Re^n$, unconstrained input control $u \in \Re^m$ and $w \in \Re^s$ denotes the exogenous input vector. The initial condition on the state vector x_0 is prescribed.

By using the dynamic programming approach[3], when the Isaacs condition[156] of interchanging the min max operation holds, the finite-time HJI equation for the above problem is given as,

$$-\frac{\partial J^*}{\partial t} = \underbrace{\min}_{u}\ \underbrace{\max}_{w}\left[H(x,u,w,t)\right] = \underbrace{\max}_{w}\ \underbrace{\min}_{u}\left[H(x,u,w,t)\right] \tag{5.4}$$

with the boundary condition

$$J^*(x(t_f),t_f) = \phi(x(t_f)); \tag{5.5}$$

---

[1] The Hamiltonian, H should be separable in u and w. Because of this condition, the operation of maxminimization or minmaximization can be interchanged. That is also a condition to obtain the game-theoretic saddle point. The detailed explanation is given by Bryson and Ho[64], Chap. 9, pg. 276.

where $x(t_f) = \{x \mid \psi(x) = \psi_f\}$, J^* is the smooth optimal return function (the cost-to-go) which satisfies all necessary and sufficient conditions for the optimality and H is defined as,

$$H(x,u,w,t) = L(x,u,w,t) + \left(\frac{\partial J^*}{\partial x}\right)^T f(x,u,w,t) \tag{5.6}$$

Series Solution Methodology

In the classical optimal control theory, first order necessary conditions[64] (also known as Euler-Lagrange equations) are given as,

$$\dot{x} = \frac{\partial H}{\partial \lambda}; \dot{\lambda} = -\frac{\partial H}{\partial x} \tag{5.7}$$

$$H_u = 0; H_w = 0 \tag{5.8}$$

and the transversality condition[65]:

$$\lambda(t_f) = \frac{\partial \phi(x(t_f))}{\partial x(t_f)} + \left(\frac{\partial \psi(x(t_f))}{\partial x(t_f)}\right)^T \nu \tag{5.9}$$

where λ is costate vector and ν is known as terminal Lagrange multiplier associated with the terminal constraint given in Eq. (5.3). For the given initial condition, the open-loop solution can be obtained by solving the two-point-boundary-value problem formed by Eqs. (5.7)-(5.9). Just like the procedure developed for the HJB equation, the HJI equation can be treated in a similar way. The key aspects of the methodology can be summarized as follows,

1. Assume a higher order polynomial series representation of the cost-to-go, J^* in terms of the state variables x, together with the terminal Lagrange multipliers ν. Each term in the series is weighted by an unknown time-dependent gain.

2. Substitute the expansion of J^* into Eq. (4) and collect the coefficients of various orders of x and ν to obtain all the gain differential equations. Eqs (5.5) and (5.9) can be utilized to find the boundary conditions to integrate the gain differential equations backwards in time. The gains can be stored for the forward propagation of the dynamical system. Now, at all stored points, the explicit analytical form of $J^*(x,\nu)$ can be known from its expression.

3. Calculate ν at each time step by using the vector series reversion[66] of Eq. (2.14) and obtain the optimal control and the exogenous output by using Eq.(5.8).

5.2 Application-I: Nonlinear Pursuit Evasion Games

New non-linear feedback strategies can be proposed for pursuit and evasion scenarios involving space assets. Pursuit-evasion games are governed by the HJI equation. The innovation is that non-linear feedback solutions to non-linear finite-time capture or circumnavigation problems are synthesized.

Consider two spacecraft in neighboring orbits. The orbit elements of the two vehicles differ slightly. One vehicle is labeled P whereas the other is labeled E. The

game scenario is for spacecraft P to efficiently approach spacecraft E, while E wishes to escape or delay interception. The governing equation of each vehicle is the point-mass model in an inverse-square law gravity field subject to radial and transverse controls.

The equations of motion of the pursuer (P) and the evader (E) spacecraft are given as follows,

$$\begin{aligned}
\dot{r}_P &= w_P ; \dot{v}_{P_T} = -\frac{w_P v_{P_T}}{r_P} + u_{2_P} ; \dot{w}_P = \frac{v_{P_T}^2}{r_P} - \frac{\mu}{r_P^2} + u_{3_P} \\
\dot{r}_E &= w_E ; \dot{v}_{E_T} = -\frac{w_E v_{E_T}}{r_E} + u_{2_E} ; \dot{w}_E = \frac{v_{E_T}^2}{r_E} - \frac{\mu}{r_E^2} + u_{3_E}
\end{aligned} \tag{5.10}$$

In Eq.(5.10), the subscript *P* denotes the pursuing vehicle whereas *E* denotes the evading vehicle. Furthermore, r denotes the radial position, v is the tangential velocity component, w denotes the radial velocity component, u_2 is the control along the tangential direction, u_3 is the control along the radial direction, and μ is the gravitational parameter in the chosen celestial frame. For the state-space representation with $x \in \Re^{6\times 1} \equiv [r_P ; v_{P_T} ; w_P ; r_E ; v_{E_T} ; w_E]$ $u \in \Re^{2\times 1} \equiv [u_{2_P} ; u_{3_P}]$ and $w \in \Re^{2\times 1} \equiv [u_{2_E} ; u_{3_E}]$, the equations of motion are given by

$$\dot{x} = f(x, u_P, u_E, t) \equiv \begin{bmatrix} x_3 \\ -\frac{x_3 x_2}{x_1} + u_{2_P} \\ \frac{x_2^2}{x_1} - \frac{\mu}{x_1^2} + u_{3_P} \\ x_6 \\ -\frac{x_6 x_5}{x_4} + u_{2_E} \\ \frac{x_5^2}{x_4} - \frac{\mu}{x_4^2} + u_{3_E} \end{bmatrix} \tag{5.11}$$

The objective is to obtain feedback control laws, which minimize the weighted square of the terminal-miss at a fixed final time. In general, the game becomes more meaningful if the control bounds are prescribed for both the pursuer and the evader. Basically, to relax the control inequality constraints, the cost function is given in two parts as,

$$J = \varphi + \frac{1}{2}\int_{t_0}^{t_f} (R_{2_P} u_{2_P}^2 + R_{3_P} u_{3_P}^2 - R_{2_E} u_{2_E}^2 - R_{3_E} u_{3_E}^2) dt \tag{5.12}$$

where,

$$\varphi = \frac{1}{2}\{S_1[r_P(t_f) - r_E(t_f)]^2 + S_2[v_{P_T}(t_f) - v_{E_T}(t_f)]^2 + S_3[w_P(t_f) - w_E(t_f)]^2\} \tag{5.13}$$

and $R_{2_P}, R_{3_P}, R_{2_E}, R_{3_E}, S_1, S_2$ and S_3 are positive weights.

The soft constraint, φ defines the weighted terminal miss whereas its integrand represents the bounded energies[64] of each of the participants in the game. Now, the optimal control problem can be defined as minimizing J subject to Eq. (5.11) with the prescribed initial conditions.

Solution Procedure

We first define the Hamiltonian, *H* as,

$$H = \frac{1}{2}(R_{2_P}u_{2_P}^2 + R_{3_P}u_{3_P}^2 - R_{2_E}u_{2_E}^2 - R_{3_E}u_{3_E}^2) + \lambda_{P_1}(x_3) + \lambda_{P_2}(-\frac{x_2x_3}{x_1} + u_{2_P}) + \\ + \lambda_{P_3}(\frac{x_2^2}{x_1} - \frac{\mu}{x_1^2} + u_{3_P}) + \lambda_{E_1}(x_6) + \lambda_{E_2}(-\frac{x_5x_6}{x_4} + u_{2_E}) + \lambda_{E_3}(\frac{x_5^2}{x_4} - \frac{\mu}{x_4^2} + u_{3_E}) \tag{5.14}$$

where λ_{P_i} and λ_{E_i} $(i = 1,2,3)$ are the costates of *P* and *E*, respectively.

The first order necessary conditions for optimality can be derived by using Eqs (5.7)-(5.9). Since the proposed series solution method assumes a polynomial form of the plant dynamics, the Hamiltonian can be expanded about the initial circular orbit of *E*. For circular orbit, the reference control and costates are zero. Since the optimal control problem is posed with a soft terminal constraint, the series expansion of J^* can be assumed to be a function of only δx defined as follows:

$$\delta x = \begin{bmatrix} x_1 - R_E \\ x_2 - \sqrt{\frac{\mu}{R_E}} \\ x_3 - 0 \\ x_4 - R_E \\ x_5 - \sqrt{\frac{\mu}{R_E}} \\ x_6 - 0 \end{bmatrix} \tag{5.15}$$

Applying necessary and sufficient conditions for saddle-point equilibrium strategies, the HJB equation becomes the HJI equation[5] given as follows,

$$\frac{\partial J^*}{\partial t} = -\underbrace{\min}_{(u_{2/3_P})}\ \underbrace{\max}_{(u_{2/3_E})}\{H(\delta x, u_{2_{P/E}}, u_{3_{P/E}}, t)\}, J(\delta x(t_f), t_f) = \varphi, \tag{5.16}$$

The game theoretic saddle point conditions[64] given as Eq. (5.8) yield the following feedback control laws,

$$\begin{aligned} u_{2_P} &= -\frac{1}{R_{2_P}}\left(\frac{\partial J^*}{\partial \delta x_2}\right); u_{3_P} = -\frac{1}{R_{3_P}}\left(\frac{\partial J^*}{\partial \delta x_3}\right) \\ u_{2_E} &= \frac{1}{R_{2_E}}\left(\frac{\partial J^*}{\partial \delta x_5}\right); u_{3_E} = \frac{1}{R_{3_E}}\left(\frac{\partial J^*}{\partial \delta x_6}\right) \end{aligned} \tag{5.17}$$

In order to solve Eq. (5.16)along with Eq.(5.15), a fourth order power series is assumed with the time-dependent gains as,

$$J^* = \sum_{k=0}^{4} f_k(\delta x_i, t), i = 1,..,6 \tag{5.18}$$

where *k* is the degree of the homogeneous polynomial expansions of its arguments. The order of nonlinear feedback is one less than the order of the cost-to-go expansion. Substituting Eq.(5.18) in Eq. (5.16) and following the procedure discussed above, results in a third order feedback law which is applied to the full nonlinear system in Eq.(5.10).

Numerical Example and Discussion

To demonstrate the method's applicability to nonlinear pursuit-evasion games, results for one example are illustrated, which involves two spacecraft in orbit about the Earth. The prescribed initial conditions are,

$$\begin{aligned} r_P(t_0) &= 6.60R; v_P(t_0) = \sqrt{1/(6.60)}; w_P(t_0) = 0 \\ r_E(t_0) &= 6.57R; v_E(t_0) = \sqrt{1/(6.57)}; w_E(t_0) = 0; \end{aligned} \tag{5.19}$$

where R is the Earth's radius and the final time is $t_f = 3.5$ hrs, which is equivalent to 15.5378 TU (For this problem, 1TU=810.9259 seconds). Other parameters for the simulations are specified as follows,

$$R_{2_P} = 0.27e-4, R_{3_P} = \frac{5}{9}R_{2_P}, R_{2_E} = 0.4, R_{3_E} = \frac{1}{2}R_{2_E}; S_1 = 1.2e-4; S_2 = 1e-5; \\ S_3 = 0.002 \tag{5.20}$$

The general expansion of the cost-to-go up to fourth order contains 210 time dependent gains. The gain differential equations are obtained by using the symbolic tool box "Maple®". Matlab® 7.0 was utilized to integrate them numerically by using a fourth-fifth order Runge-Kutta method (ode45); values for each gain were stored at 300 discrete points and supplied for integrating the nonlinear dynamics forward in time. The required feedback control input were computed with the use of Eq. (5.17).

Figure 5.1 illustrates simulation results performed for several pursuer initial conditions. To check the performance of the synthesized feedback control, the time-dependent feedback gains for all simulations were computed for one set of initial conditions (corresponding to the heavy solid line) and then used for the other sets. The results are reflected in the cost-to-go plot shown as Fig. 5.2. The terminal miss distance, which is specified as a soft terminal constraint, changed slightly for the different initial conditions, but the range in the terminal miss stayed within 100 m. The key point is that the nonlinear feedback approach eliminates the need to re-compute gains for differing initial conditions. On the other hand, it also shows the robustness of the nonlinear controller.

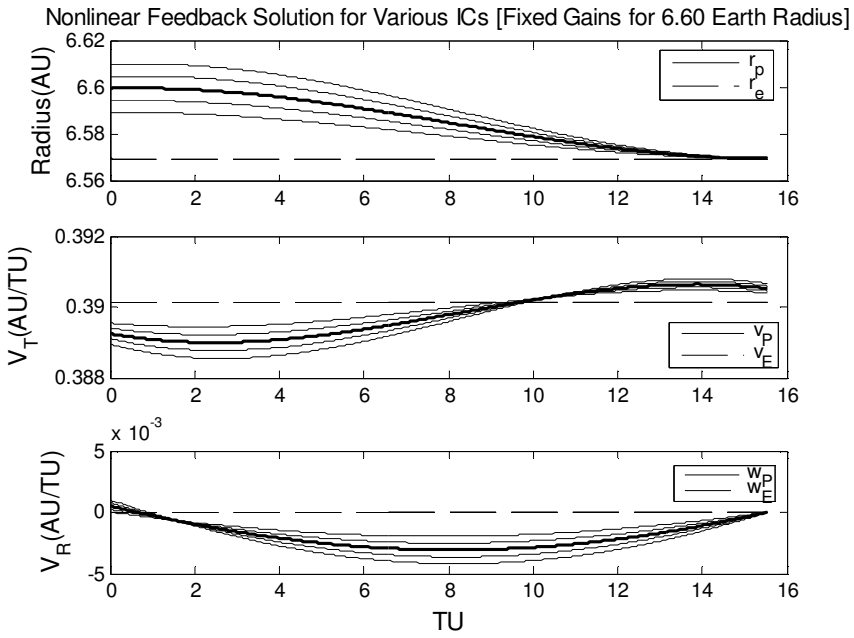


Fig. 5.1 States of the Pursuer and the Evader

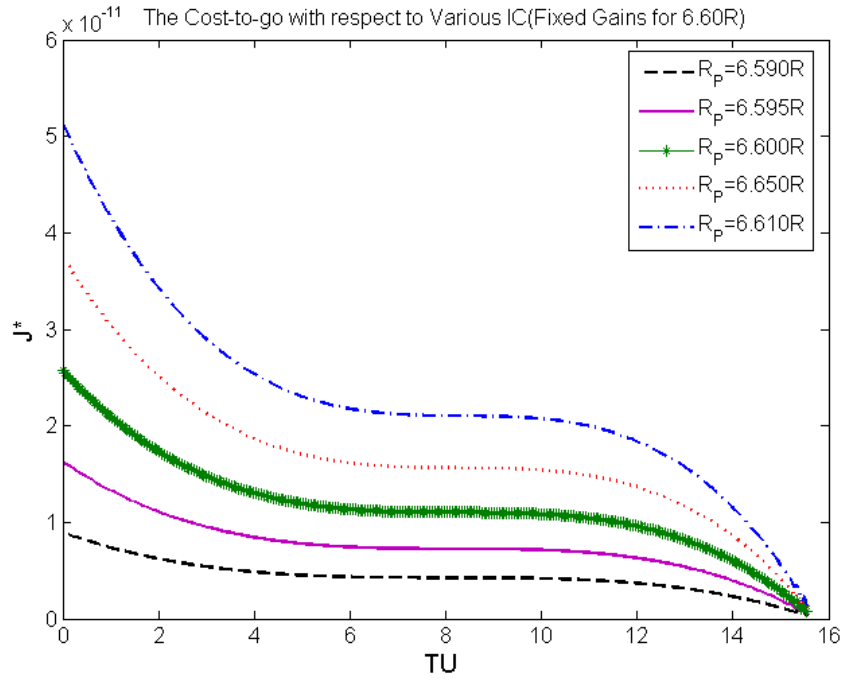


Fig. 5.2 The Cost-to-go with Respect to Various ICs with the Fixed Gain

5.3 Application-II: Finite-time H_∞ Feedback Controllers for Terminally Constrained Nonlinear Systems

The second major HJI formulation is encountered in the field of robust optimal control theory. With the game-theoretic interpretation, H_2 / H_∞ problems can also be cast as minimax problems, which, as previously discussed, lead to the HJI equation formulations[64]. It is shown in Ref. [45] that H_2 / H_∞ control laws can also be viewed as a zero-sum two player differential game. Theoretically, it can be stated that in the given multi-input, multi-output system, the performance of the designed controller will be effective if it can facilitate a stabilizing feedback control input with the attenuation of the worst known disturbances. H_∞ control theory provides the practical construction of such controllers. Detailed treatment of the formulation and solution of the H_∞ closed-loop problem can be found in references [49, 55].

The next section presents a brief overview of the HJI formulation associated with the robust optimal control problems. Application of the SSM to a missile guidance example is treated and the result is compared with the open-loop solution for several cases.

Problem Formulation

Consider a nonlinear state-space model,

$$\begin{aligned} \dot{x} &= f(x) + g_1(x)w + g_2(x)u \\ z &= h_1(x) + k_{12}(x)u \end{aligned} \tag{5.21}$$

where $x \in \Re^{n\times 1}, u \in \Re^{m_1\times 1}, w \in \Re^{m_2\times 1}$ represent state, control and disturbance vector of appropriate dimensions, respectively.

In the optimal control sense, we want to minimize the following performance metric,

$$J = \frac{1}{2}\int_{t_0}^{t_f} [\|h_1(x) + k_{12}(x)u\|^2 - \gamma^2 \|w\|^2]dt \tag{5.22}$$

Subject to Eq. (5.21) with prescribed initial and boundary conditions

In order to simplify the design of the controller, the following conditions are also applied,

$$\begin{aligned} h_1^T(x)k_{12}(x) &= 0 \\ k_{12}^T(x)k_{12}(x) &= R_2 > 0 \end{aligned} \tag{5.23}$$

The attenuation factor, $0 < \gamma$ can be chosen such that the response z should satisfy the following condition,

$$\int_{t_0}^{t_f} z^T(\tau)z(\tau)d\tau \le \gamma^2 \int_{t_0}^{t_f} w^T(\tau)w(\tau)d\tau \tag{5.24}$$

Solution Procedure

The H_∞ problem boils down to the standard Bolza form optimal control problem. By using the dynamic programming approach, we obtain the HJI equation as shown below:

$$-\frac{\partial J^*}{\partial t} = \underbrace{\min}_{u}\ \underbrace{\max}_{w}\left(\frac{1}{2}[\|h_1(x) + k_{12}(x)u\|^2 - \gamma^2\|w\|^2] + \left(\frac{\partial J^*}{\partial x}\right)^T f(x) + g_1(x)w + g_2(x)u\right) \tag{5.25}$$

The necessary conditions for optimality result in the feedback form for control input, u and the attenuated disturbance, w:

$$u = -R_2^{-1} g_2^T(x)\left(\frac{\partial J^*}{\partial x}\right)$$
$$w = \frac{1}{\gamma^2} g_1^T \left(\frac{\partial J^*}{\partial x}\right) \tag{5.26}$$

Feedback control and disturbance forms can be obtained via SSM.

5.4 Short Pitch Dynamics Model of a Missile

An application of SSM to the control of the short pitch dynamics model of a missile [41] is presented in this section. The nonlinear pitch dynamics of missile is given as,

$$\dot{\alpha} = Q + \frac{qS}{mV_m} C_z(\alpha, M, \delta_e)$$
$$\dot{Q} = \frac{qSc}{I_{yy}} C_m(\alpha, M, \delta_e) + w \tag{5.27}$$
$$\dot{\alpha}_i = \alpha$$

where,

$$C_z = -0.5052\alpha + 0.0429 + (0.1230\alpha - 0.0191)M + 0.09\delta_e$$
$$C_m = -0.0055\alpha^3 + 0.2131\alpha^2 - 2.7419\alpha - 0.0381 + \tag{5.28}$$
$$+(0.0014\alpha^3 - 0.0623\alpha^2 + 0.8715\alpha - 0.4041)M - 0.675\delta_e$$

and α, Q and α_i are the angle of attack, pitch rate and integral control, respectively. The state vector is $x \equiv [\alpha, Q, \alpha_i]$, control input is $u \equiv \delta_e$, and the exogenous disturbance is

denoted by *w*. Other parameters to define the metric and for use in the equations of motion (5.27) are given as,

$$\begin{aligned} &k_1 = 32; k_2 = 1; k_3 = 520; \rho_1 = 21; \rho_2 = 7 \\ &M = 2.75 \end{aligned} \tag{5.29}$$

Before we proceed to deal with the HJI equation, the effect of the attenuation factor, γ which essentially characterizes the effect of disturbances in the performance index, is studied. Figure 5.3 shows the open-loop solutions for the rest-to-rest, fixed final time maneuvers for various values of γ. As γ is increased, the midcourse trajectories show lesser midcourse deviations from the boundary conditions. The design for a smaller value of attenuation factor can handle a larger disturbance. Finding the best value of γ depends on the problem and the degree of nonlinearity. Based on the design requirements, the typical value of γ can be selected within the range of [1, 20]. Values beyond this range lead to conjugate points in the solutions to the OCP.

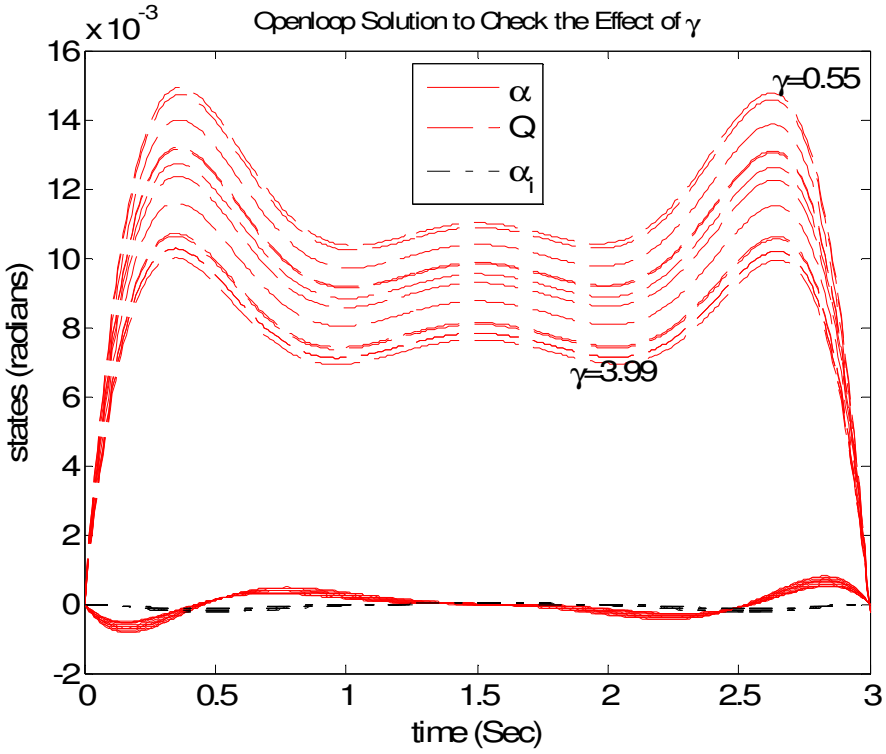


Fig. 5.3 The Effect of the Attenuation Factor

The series solution procedure described above can be followed step by step to obtain the solution of Eq. (5.25) for the missile problem. A fourth-order expansion of J^* is assumed in x and ν. All gains are stored at 100 discrete time points. During the forward integration of the given dynamics, the value of ν is computed by using a second-order vector series reversion process.

Two different cases are considered for presenting the numerical simulations:

CASE A: Rest-to-rest Missile Maneuver in Finite-time

Initial and terminal conditions are given as;

$$\begin{aligned}&\alpha(t_0)=0;\alpha(t_f)=0;Q(t_0)=0;Q(t_f)=0\\&\alpha_i(t_0)=0;\alpha_i(t_f)=0;\gamma=3.99;t_0=0;t_f=3\end{aligned} \tag{5.30}$$

CASE B: Non-zero Boundary Conditions

Initial and boundary conditions are given as;

$$\begin{aligned}&\alpha(t_0)=0.06;\alpha(t_f)=-0.2054;Q(t_0)=0.06;Q(t_f)=-0.0239\\&\alpha_i(t_0)=0.06;\alpha_i(t_f)=-0.0068;\gamma=3.1241545\ \ note:angles \text{ are in } radians\end{aligned} \tag{5.31}$$

The state trajectories for case A are shown in Figs. 5.4-5.6. Open-loop solutions are obtained by using shooting methods[68]. The control and the disturbance profile are plotted in Fig. 5.7. The bold line shows the third-order feedback solution whereas the dashed line is used to emphasize the respective open-loop solution. The plots of the state trajectories for both open-loop and feedback cases show that the third-order controller performs well to satisfy the terminal constraint. It is noted that for the range of [0.1, 1], a linear control results in a 20-50% error, clearly showing the importance of the nonlinear control law. Similar observations can be made for case B, which is simulated for non-zero initial and final conditions [Figs. 5.8-5.10]. For fixed set of initial and terminal conditions, many optimal solutions can be constructed as the value of γ is varied. Finally, it is also noted that the SSM presented is noniterative, whereas the open-loop procedure, regardless of the method utilized, is inherently iterative.

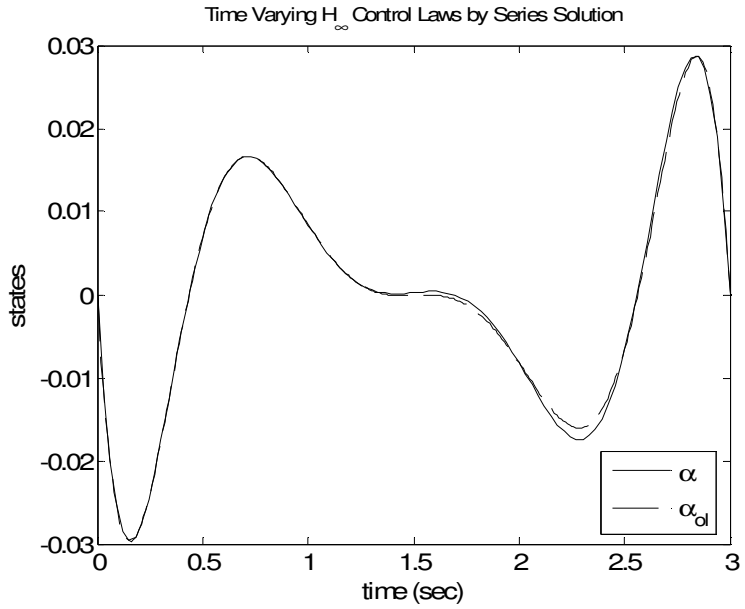


Fig.5.4 Case A: Open-loop and Feedback Solutions for the Angle of Attack for Zero Initial and Final Conditions

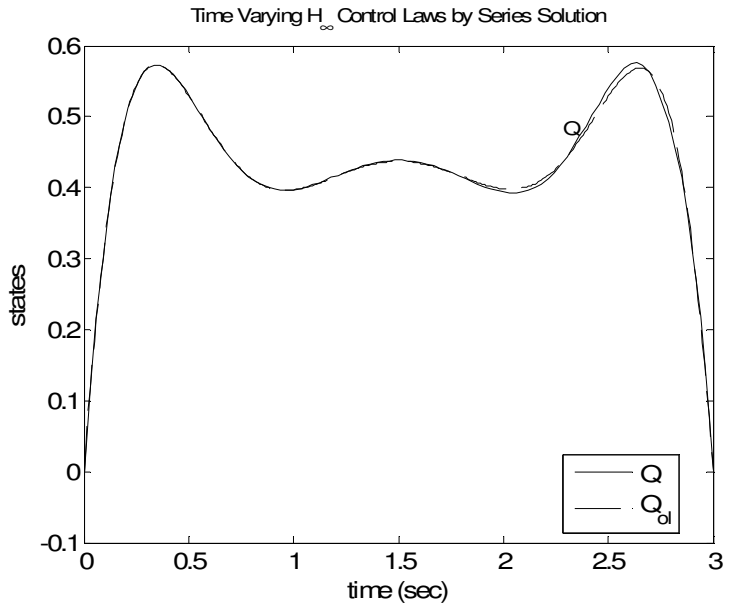


Fig. 5.5 Case A: Open-loop and Feedback Solutions for the Pitch Rate for Zero Initial and Final Conditions

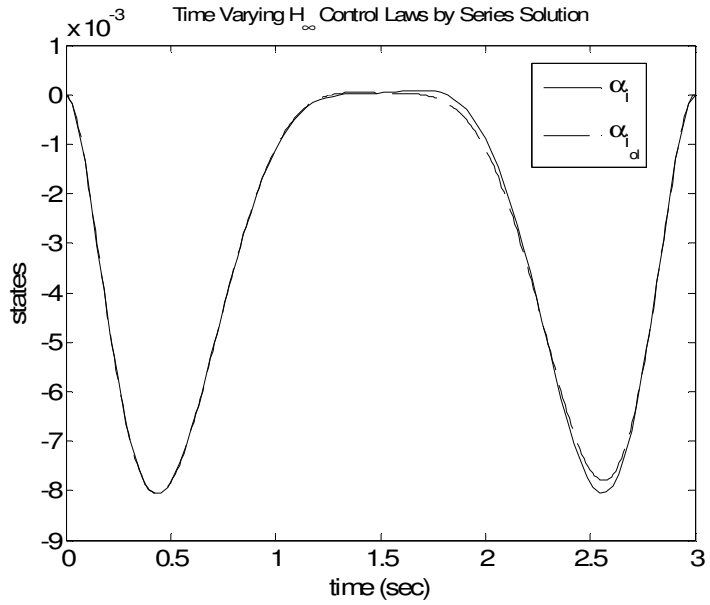


Fig. 5.6 Case A: Open-loop and Feedback Solutions for the Integral Control for Zero Initial and Final Conditions

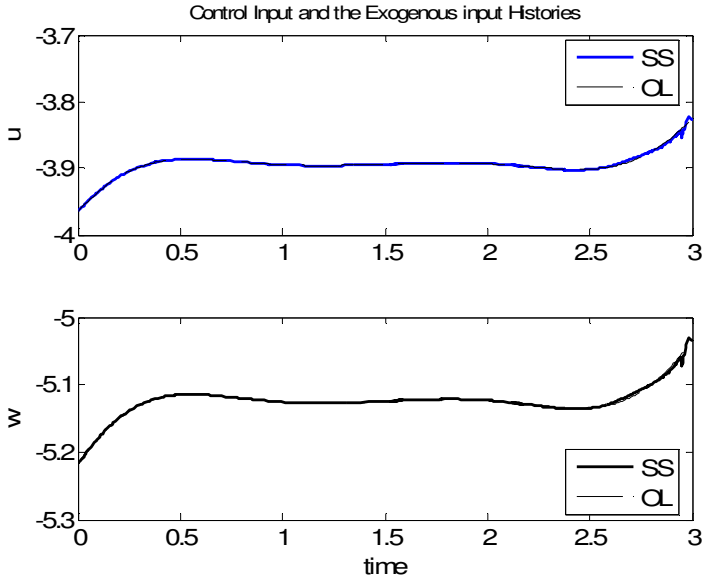


Fig.5.7 Case A: Open-loop and Feedback Solutions for the Control Input and the Disturbance

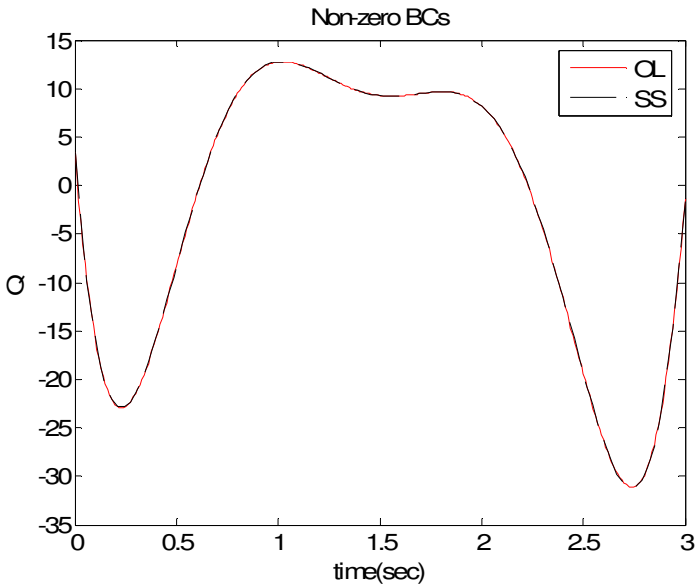


Fig. 5.8 Case B: A Comparative Study by Using Open-loop and Feedback Solutions for the Pitch Rate for Nonzero Initial and Final Conditions

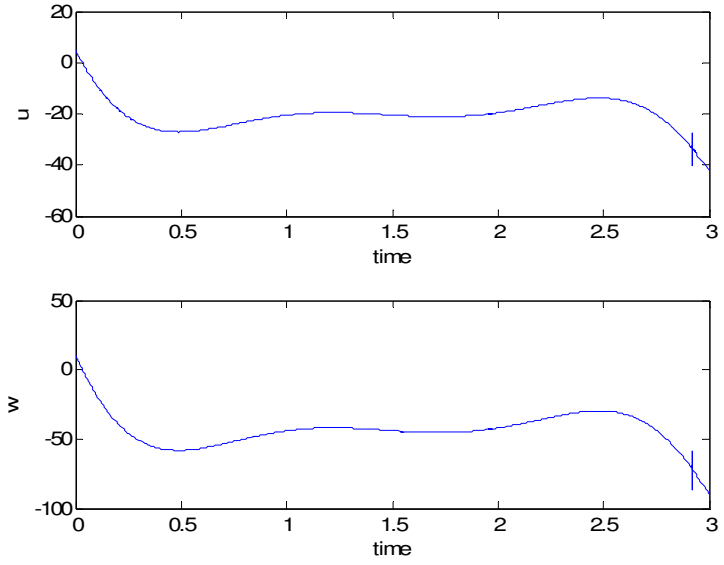


Fig 5.9 Case B: Required Control Effort (u) and the Worst Disturbance (w)

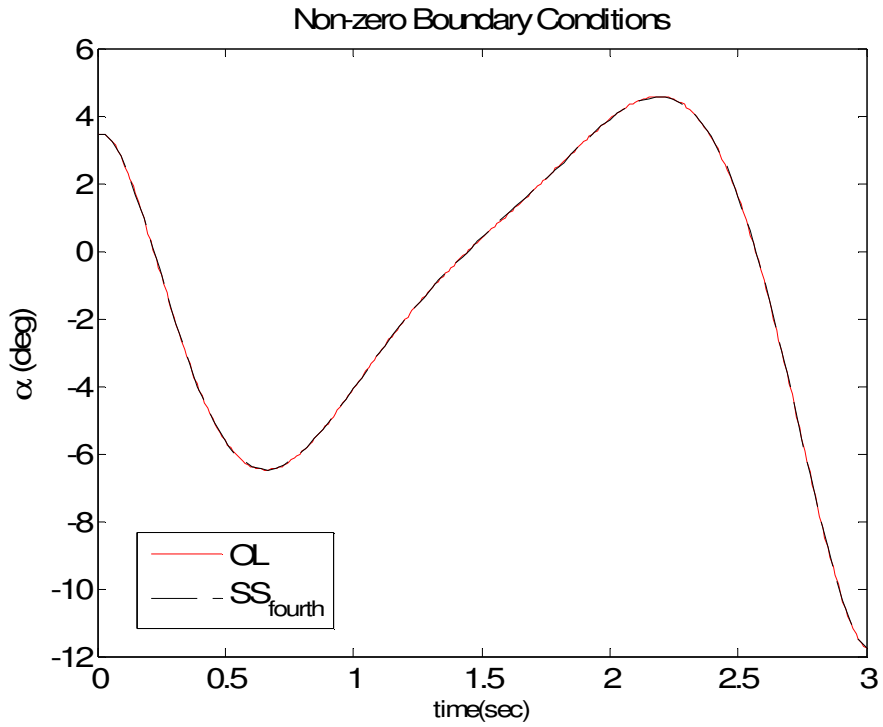


Fig. 5.10 Case B: A Comparative Study by Using Open-loop and Feedback Solutions for the Angle of Attack for Nonzero Initial and Final Conditions

CHAPTER VI

OPTIMAL STRATEGIES IN COOPERATIVE GAMES WITH TERMINAL PAYOFF

This chapter highlights two further applications of SSM to problems of dynamic game theory. SSM is applied to the Pareto optimal, finite-time, nonlinear game, subject to hard terminal constraints. The two-player problem is formulated as a time-dependent HJI equation and an optimal feedback solution is obtained. Several numerical examples are included to develop the Pareto frontier for the general quadratic payoff function with a nonlinear co-operation model and terminal constraints. The second application involves a structured cooperation approach among the players by the introduction of a linear cooperation equation. SSM is also applied to the multiplayer linear quadratic differential game problems. In order to focus the theoretical development, some simple, but concept-based numerical and analytical solutions are discussed to illustrate the desired benefits in each player's payoff function.

So far, the known pursuit-evasion techniques have been applied to soft constraint problems only; direct treatment of hard terminal constraints is unavailable in the literature. This chapter treats linear and nonlinear cooperation models and is presented in two parts. In the first part, a general methodology is developed to analyze the nonlinear deterministic continuous differential games with both soft and the hard terminal constraints. The result is an optimal feedback strategy for exhibiting the Pareto optimal

solution for a two-player nonlinear cooperative game. The methodology developed in this chapter is not limited to the cooperative case only; the same approach can be implemented to any finite horizon zero or nonzero sum game with hard terminal constraints. The second part of this chapter describes a new linear cooperative model, named the z-model. Also, a two pursuers-one evader capture game is studied numerically. The focus of this example is on the establishment and investigation of a cooperation model among the pursuers, which allows capture to occur in a more efficient manner than if the pursuers did not cooperate.

6.1 Overview of Differential Games

Pursuit-evasion differential games can be challenging optimization problems in which the pursuers influence their achievements with the opposing intents of the evaders. Such conflicting situations can be realized in various fields, including space applications, robust control, and management science or macro economics. Practical applications[69-71] include formation flying, optimum task allocation, air-traffic control, uncertainty modeling, circumnavigation and visual identification, collision avoidance designs in transport systems, and debt regulation[72]. The theory of differential games provides optimal control strategies for problems involving dynamic systems and algebraic constraints.

Once the optimization problem is cast as a game, the non-trivial part of the game is to define certain rules and specify the rules that all players must comply with until the

game terminates. Some of those rules could be strategies for cooperation or non-cooperation[8] among the group of pursuers or evaders. As mentioned above, there are many situations where the main task is to accomplish a capture or to attain a certain benefit in any players' payoff function, such as, the time of capture of all evaders, consumed fuel, terminal miss, or the exact satisfaction of the terminal constraints. Some critical scenarios, however, could be encountered if the evader is more powerful or the individual pursuers are inefficient. To resolve such cases, the establishment of cooperation among players can be of great help. Cooperation can be defined in different contexts: for instance, one may see cooperation among some players or one can define it to the need-based circumstances.

The theory of differential games has seen many advances since its inception. Although many elegant aspects of the two players, zero-sum game[60-61] have been addressed over the years, there are still some research questions to be answered. So far, even for problems of low dimensions, finding real-time optimal feedback strategies for nonlinear games is a formidable task. One possible approach is to follow the well-established formulation[7] of the HJI equation. In general, due to the known inherent difficulties in solving this nonlinear partial differential equation, the solution of the finite-horizon problem with hard terminal constraints has not yet been addressed. Many facets of cooperative linear quadratic Pareto games have been widely studied. Pareto optimality[9] is concerned with the best joint decisions among players. However, currently there is no known extension towards the Pareto solutions of nonlinear cooperative games. Continuing with two-player, zero-sum games with Nash solutions[8], we extend

the developed SSM to investigate nonlinear games. Much of the theory for two player games involving a single pursuer and a single evader was established by Isaacs[6-7] and Bryson & Ho[64]. These researchers defined and laid out important definitions, formulations, and assumptions. Even for linear quadratic differential games, significantly less development, however, has been completed for the cooperative multiplayer situation.

6.2 Cooperative Dynamic Games

In this section, we discuss the theoretical development of finite-time nonlinear differential games with terminal constraints. Before we formulate the exact problem, in general, there are following assumptions to be taken into account,

1. The players can communicate and work with the desired level of cooperation.
2. All information pertaining to the individual's performance index and dynamics is known to all players throughout the game.
3. The applied control is affine and unbounded. Furthermore, all players are able to apply it with the known information on states.
4. Each player has a quadratic cost and optimal feedback strategies are considered.

Basically, the cooperation dynamics is defined by a state vector $x(t) \in \Re^{n\times 1}$ which is influenced by all players. For an *m* player game, the state dynamics is given as,

$$\dot{x} = A_i x^i + B_j u_j,\ i = 0,1,2,...,p;\ j = 1,2,...,m \tag{6.1}$$

where $(\dot{}) \equiv d/dt$, A_i is a higher order tensor if $i \geq 2$ and u_j is control vector of the j^{th} player. Other matrices are taken of appropriate dimensions.

Every player would like to attain the desired task with the minimum effort; however, for all players, there are multiple solutions to the problem when cooperation is established among them. Searching for the undominated control strategy for each player within the given control domain underlies the Pareto optimal concept[9] :

Let u_i^* be the optimal control for the i^{th} player and J_i be the cost. The Pareto optimal solution is defined by the inequality,

$$J_i(u_1, u_2, ..., u_m) \leq J_i(u_1^*, u_2^*, ..., u_m^*); i = 1, ..., m \tag{6.2}$$

if there exists at least one strict inequality in Eq. (6.2), which has no solution for all feasible values of control.

A set of all Pareto solutions defines the Pareto frontier. For two players, Fig. 6.1 illustrates a Pareto optimal solution using '*' on the solid curve, a Pareto frontier. The cost region above the curve is available to both players.

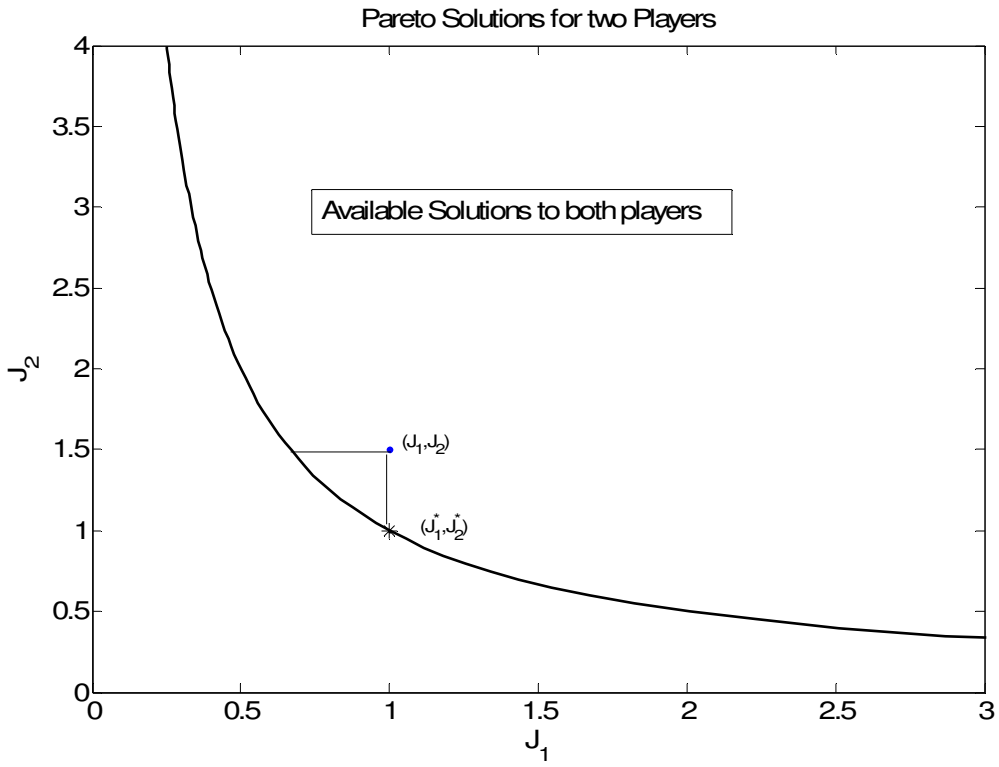


Fig. 6.1 Pareto Solutions for Two players

Our objective is to obtain a feedback solution to cooperative dynamic games with a terminal constraint.

Problem Formulation

Consider the finite horizon *i*-players differential game, in which all players are participating to minimize the following quadratic cost functions,

$$J_i = \varphi_i(T) + \frac{1}{2}\int_0^T (x^T Q_i x + R_i u_i^2 + C_i)dt; Q_i \geq 0, R_i > 0 \tag{6.3}$$

subject to

$$\psi \equiv k_1 x(T) + k_2 \dot{x}(T) + k_3 x^2(T) + k_4 \dot{x}^2(T) + k_5 x(T)\dot{x}(T) = \psi_f \tag{6.4}$$

and the 2 dimensional cooperation dynamics described as,

$$\begin{aligned} &\ddot{x} = \sum_{j=0}^{n} m_j x^j + b_i u_i; \\ &x(0), \dot{x}(0) \text{ are prescribed} \end{aligned} \tag{6.5}$$

where x is a dynamic variable that can be influenced by both players, n denotes the order of the nonlinear system and m_j are the system's parameters.

Solution Procedure

Possibly, the objectives of all or some of the players might be conflicting. To obtain all possible cooperative solutions for the optimal control problem defined by Eq. (6.3)-(6.5), a new parameter α_i can be introduced to obtain a combined cost function as follows,

$$\begin{aligned} &\Im = \sum \alpha_i J_i; \\ &\sum \alpha_i = 1; 0 < \alpha_i < 1 \end{aligned} \tag{6.6}$$

As the value of α is varied from 0 to 1, a different Pareto solution is obtained. For linear systems, Lancaster et. al[54] has proved the Pareto frontier will be a smooth function of α for finite horizon problems. For the sake of simplicity, we also assume a linear combination of all individual cost functions as shown in Eq. (6.6).

The following steps can be applied to solve for the optimal feedback strategy,

1. Construct the Hamiltonian, H and utilize the necessary conditions for optimality[64] to derive the HJI equation for the cooperative game problem. For ease of exposition, we derive it for the 2-player case as,

$$\frac{\partial \Im^*}{\partial t} = -\underbrace{\min}_{u_1,u_2}\left[\frac{1}{2}\sum_{i=1}^{2}(x^T Q_1 x + R_i u_i^2 + C_i) + \frac{\partial \Im^*}{\partial x}(\dot{x}) + \frac{\partial \Im^*}{\partial \dot{x}}(\sum_{j=0}^{n}(m_j x^j) + b_1 u_1 + b_2 u_2)\right] \tag{6.7}$$

with the boundary condition

$$\Im^*(x(t_f), t_f) = \varphi(t_f); \quad \text{where } x(t_f) \text{ satisfies the terminal constraint } \psi = \psi_f \tag{6.8}$$

2. The optimal feedback controls for each player can be obtained from the following equation,

$$u_i = -\frac{b_i}{R_i}\left(\frac{\partial \Im^*}{\partial \dot{x}}\right); i = 1, 2 \tag{6.9}$$

3. Solve the HJI equation for various values of α_i to obtain the individual cost function and draw the Pareto frontier.

6.3 Numerical Examples

To demonstrate the methodology, we consider a 2 players game with a variety of nonlinear cooperation models and soft and hard terminal constraints. Three different cases are considered to exhibit the performance of the proposed algorithm:

Case A: Stable Nonlinear Cooperation Model with A Soft Terminal Constraint

$$J_i = \frac{1}{2}\int_0^{t_f} R_i u_i^2; i = 1,2$$

subject to

$$\ddot{x} = -x - 0.6x^3 + u_1 + 2u_2; x(0) = [.1; 0.2]; \tag{6.10}$$

$$\varphi \equiv \frac{1}{2}[s_1(x(t_f) - 0.5)^2 + s_2(\dot{x}(t_f) - 1)^2]; R_1 = 1; R_2 = 1.5; t_f = 5, s_1 = 0.6; s_2 = 0.8$$

Case B: Stable Linear Cooperation Model with A Hard Terminal Constraint

$$J_i = \frac{1}{2}\int_0^{t_f} R_i u_i^2, i = 1,2$$

subject to

$$\ddot{x} = -0.5x + u_1 + u_2; x(0) = [.106; 0.212] \tag{6.11}$$

$$\psi \equiv 0.1x(t_f) + 0.1\dot{x}(t_f) = 1;$$

$$R_1 = 1; R_2 = 0.5; t_f = 5$$

Case C: Unstable Nonlinear Cooperation Model with A Nonlinear Hard Terminal Constraint

$$J_i = \frac{1}{2}\int_0^{t_f} (C_i + x^T Q_i x + R_i u_i^2) dt, i = 1,2$$

subject to (6.12)

$$\ddot{x} = -0.1 + 0.25x - 0.2x^2 + 0.7x^3 - 0.5x^4 - 0.5x^5 - 0.7x^6 + u_1 + 2u_2;$$

$$\psi \equiv 0.5x(t_f) + 0.1\dot{x}(t_f) + 0.2x(t_f)^2 + 0.1\dot{x}(t_f)^2 + 0.1x(t_f)\dot{x}(t_f) = 0.25 \tag{6.13}$$

$$R_1 = 1; R_2 = 0.5; C_1 = 0.2; C_2 = 0.4; t_f = 2;$$

$$Q_1 = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.25 \end{bmatrix}; Q_2 = \begin{bmatrix} 0.25 & -0.1 \\ -0.1 & 0.2 \end{bmatrix}; x(0) = [-0.05; 0.2] \tag{6.14}$$

In all the cases considered, a sixth-order expansion of $\mathfrak{I}^*$ is assumed in x and ν. All the gains are stored at 200 discrete time points. Due to the hard terminal constraint in cases B and C, the SSM requires the computation of the terminal Lagrange multiplier at each time step in the forward integration of the closed-loop dynamics. The value of ν is computed by using the fourth-order vector series reversion process.

Figure 6.2 demonstrates the Pareto frontier for case A in which a nonlinear stable cooperation dynamics is considered with a soft terminal constraint. The individual costs J_1 and J_2 are shown for various values of α ranging from 0 to 1. It is clearly demonstrated in Fig. 6.3 that the terminal miss, if posed with a soft constraint, varies as the value of α is changed Theoretically, α can be thought of as a cooperation factor between the conflicts of two players. By studying these plots, players can program a priori for a significant level of cooperation if they are determined to achieve the desired terminal miss.

Figure 6.4 shows the results for a linear system with a hard terminal constraint. For a linear system, there is no terminal error for any Pareto solution, unlike that for the soft constraint case.

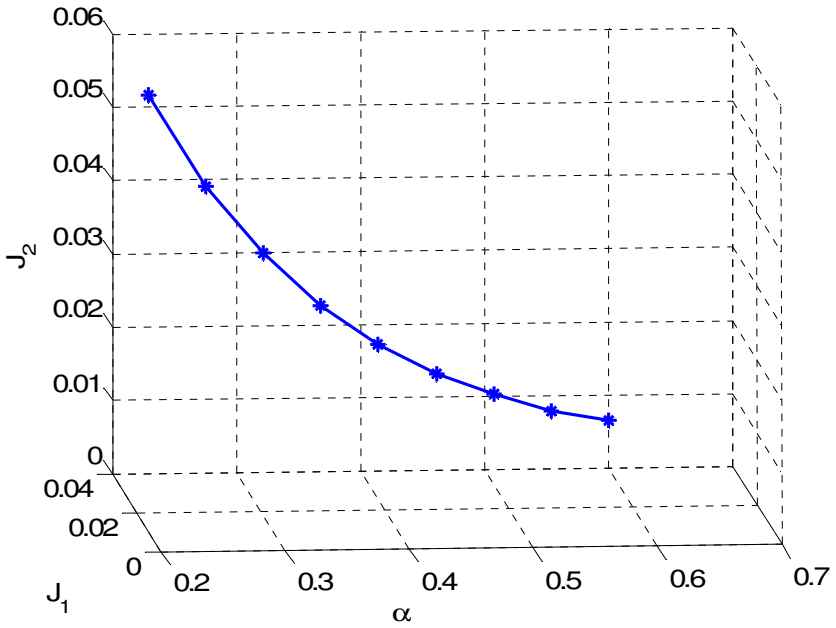


Fig. 6.2 Case A: Pareto Frontier with a Soft Constraint

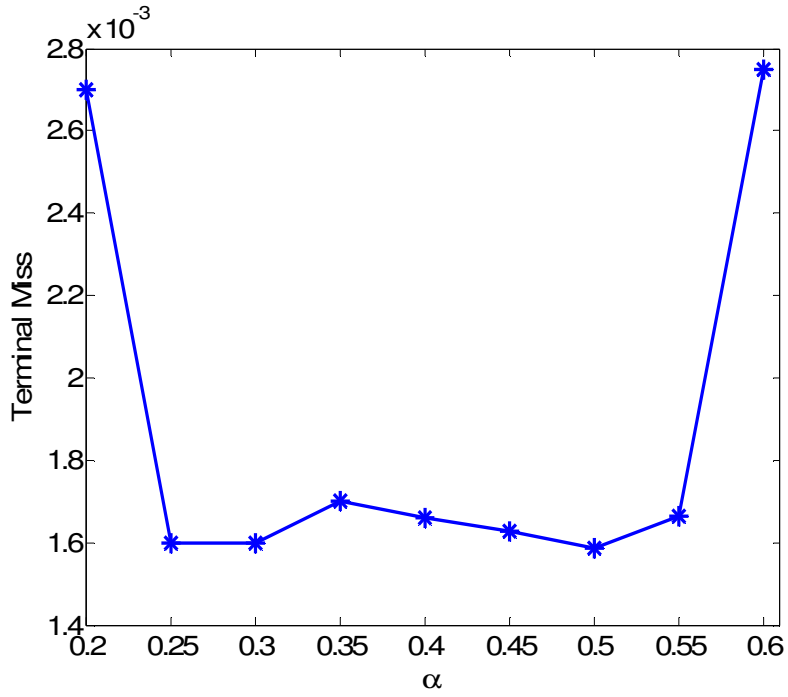


Fig. 6.3 Case A: Terminal Miss with Respect to α

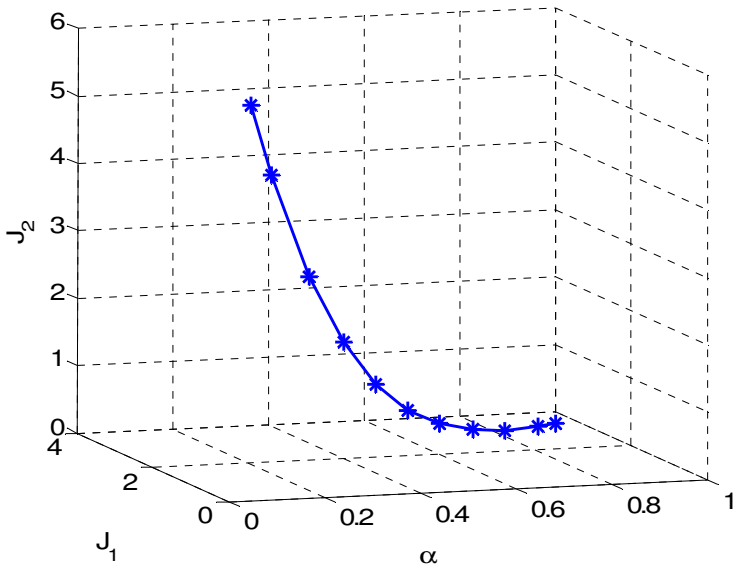


Fig. 6.4 Case B: Pareto Frontier with a Hard Terminal Constraint

The third example (Case C) addresses a two-dimensional nonlinear cooperation model which contains higher order terms up to sixth order. The nonlinear terminal constraint is defined as a hyper-surface and each player has a metric given by Eq. (6.12). In Fig. 6.5, a phase-space plot illustrates the existence of field of extremals with different values of α. The black bold line is for the lowest value of $\alpha = 0.15$ considered, for the two player cooperation whereas the dotted black line is for a high level cooperation, $\alpha = 0.6$. A terminal constraint error of 1e-4 % results due to the sixth order series expansion of $\Im^*$ and the truncation of the series reversion process. The Pareto frontier for this generalized nonlinear case is shown in Fig. 6.6. Even for this case, it is observed that, based on player's intentions, there are many available Pareto solutions

and the frontier clearly depicts the boundary of the region of the achievable minimal values for every player with known binding agreements. Computing the best situation among all Pareto solutions requires extensive investigation under the bargaining theory[62-63], which is beyond the scope of this chapter.

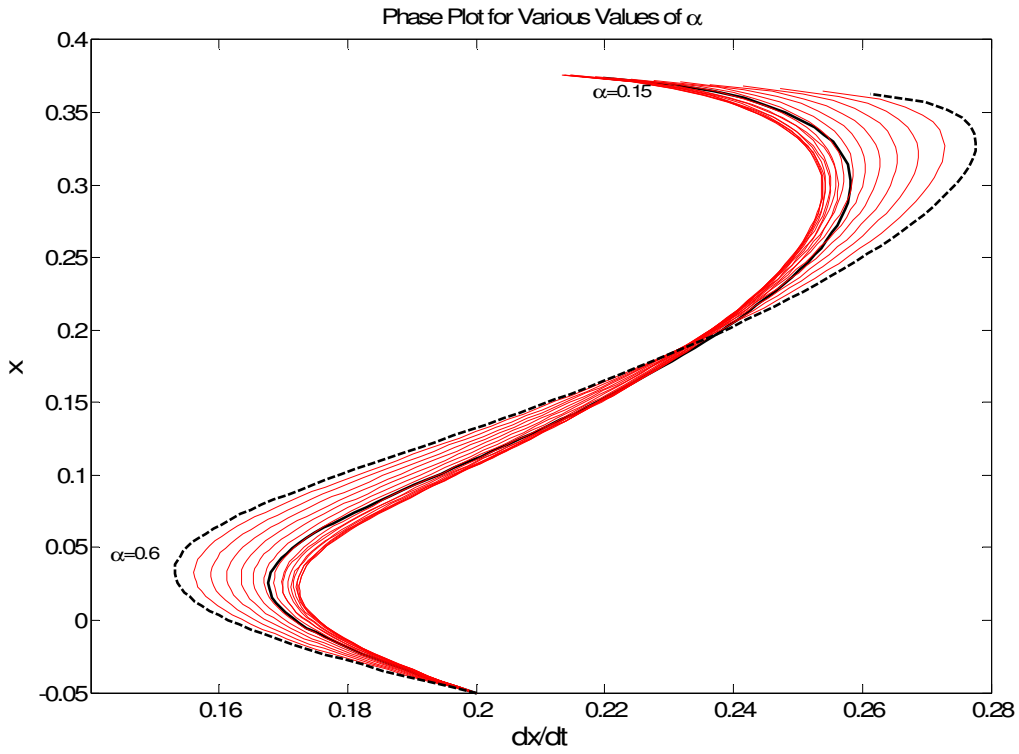


Fig. 6.5 Case C: Solution Structure of Nonlinear Cooperation

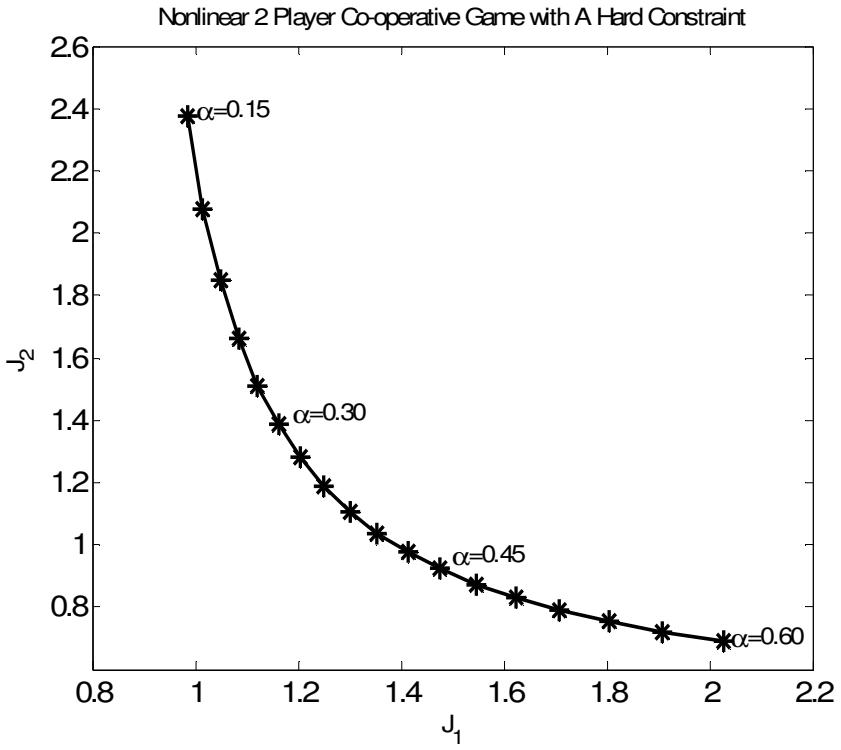


Fig. 6.6 Case C: Pareto Frontier for Highly Nonlinear Case with a Hard Constraint

6.4 Feedback Solution of Linear Quadratic Cooperative Games

The focus of this section is the establishment and investigation of a cooperation model among the pursuers that are attempting to capture a single evader. The goal of the cooperation model is to enable capture to occur in a more efficient manner than if the pursuers did not cooperate. Typically, cooperation among players is cast and studied via Pareto optimal solutions. The basis of Pareto optimal solutions is that a family of optimal solutions is generated, the differentiating factor among them being how much or how

little a player wishes to cooperate. The next section defines the cooperation model, termed the z-model.

6.5 The *z*-model

We introduce a cooperation state, which is essentially a hidden state that evolves according to the pursuer states and controls. The model parameters for the hidden state are selected so that the inclusion of the state as part of the overall system enables capture to occur in a more efficient manner than if the pursuers did not cooperate. Let's consider a multiplayer pursuit-evasion game, with *i* pursuers and *j* evaders, described as follows:

Minimize

$$J = \varphi(x_{i,j}(t_f)) + \int_{t_0}^{t_f} L(x_{i,j}, u_{i,j}, t)dt \tag{6.15}$$

subject to

$$\dot{x}_{i,j} = f_{i,j}(x_{i,j}, u_{i,j}, t); x_{i,j}(t_0) = x_{i,j_0} \tag{6.16}$$

The payoff function is defined as a soft constraint, $\varphi(x_{i,j}(t_f))$. Let the final outcome obtained by open-loop optimization be $\varphi_f(x_{i,j}(t_f))$. Since our objective is to obtain a smaller payoff value than $\varphi_f(x_{i,j}(t_f))$ which is intended using without any cooperation among *i* players, we introduce a linear z-model as,

$$\dot{z} = AX + BU$$
where, (6.17)
$$X = [z, x_1, x_2, ..., x_i]^T; U = [u_1, u_2, ..., u_i]^T$$

and A and B are the matrices of appropriate dimensions. The original pursuit-evasion problem is recast with a z-dependent quadratic term, $L_1(z)$ as,

Minimize:

$$J = \varphi(x_{i,j}(t_f)) + \int_{t_0}^{t_f} [L(x_{i,j}, u_{i,j}, t) + L_1(z)]dt \tag{6.18}$$

subject to

$$\begin{aligned} &\dot{x}_{i,j} = f_{i,j}(x_{i,j}, u_{i,j}, t); x_{i,j}(t_0) = x_{i,j_0} \\ &\dot{z} = AX + BU; z(t_0) = z_0 \end{aligned} \tag{6.19}$$

Now, choosing proper A, B and L_1 yields the desired reduction in the payoff function. In other words, this artifice allows for a means to introduce time-varying weights in the performance index. This cooperation strategy is clearly illustrated by the following analytical and numerical examples.

6.6 A Two-player Example with an Analytical Solution

A simple two-player example is constructed to investigate the z-model approach analytically. It is assumed that the evader is fixed at the origin and two pursuers desire to capture it. To investigate the benefits in the final outcome, analytically, we will obtain

the payoff function and terminal miss by using "no cooperation" as well as the "*z*-model".

The multiplayer pursuit evasion problem is posed as the following optimal control problem;

Min

$$J = \varphi + \frac{1}{2}\int_0^{t_f} (R_{P_1} u_{P_1}^2 + R_{P_2} u_{P_2}^2)dt, R_{P_{1,2}} > 0$$

subject to (6.20)

$$\dot{x}_{P_1} = u_{P_1}; \dot{x}_{P_2} = u_{P_2}; x_{P_1}(0) = x_{1_0}; x_{P_2}(0) = x_{2_0}$$

$$\varphi = \frac{1}{2}(x_{f_1}^2 + x_{f_2}^2)$$

Here, the payoff function, φ is the square of the distance at the final time. The two controls associated with the pursuers are u_{P_1} and u_{P_2}.The payoff function for the no-cooperation case, φ_{NC} can be obtained by using the necessary conditions for optimality[64],

$$\varphi_{NC} = \frac{1}{2}\left[\left(x_{1_0} - \frac{x_{f_1} t_f}{R_{P_1}}\right)^2 + \left(x_{2_0} - \frac{x_{f_2} t_f}{R_{P_2}}\right)^2\right] \tag{6.21}$$

The same problem can be solved by using the linear *z*-model defined as,

$$\dot{z} = b_1 u_{P_1} + b_2 u_{P_2}; z(0) = 0; \tag{6.22}$$

and

$$L_1(z) = \frac{1}{2} q z^2; \tag{6.23}$$

b_1, b_2 and $q \geq 0$ are constants which can be prescribed.

The payoff function with cooperation, φ_C can also be obtained analytically by using the necessary conditions as shown below:

$$\varphi_C = \frac{1}{2}\left[\left(x_{1_0} - \frac{x_{f_1} t_f}{R_{P_1}} - \beta_1\right)^2 + \left(x_{2_0} - \frac{x_{f_2} t_f}{R_{P_2}} - \beta_2\right)^2\right] \tag{6.24}$$

where,

$$\beta_m = \frac{b_m q}{R_{P_m}}\left[\frac{(b_1 x_{f_1} + b_2 x_{f_2})\{e^{\alpha t_f}(1-\alpha t_f) + e^{-\alpha t_f}(1+\alpha t_f) - 2\}}{\alpha^3 (e^{\alpha t_f} - e^{-\alpha t_f})}\right] > 0; m = 1,2 \tag{6.25}$$

$$\alpha^2 = \left(\frac{b_1^2}{R_{P_1}} + \frac{b_2^2}{R_{P_2}}\right) \tag{6.26}$$

It can be seen that β_m is a positive quantity, if x_{f_1}, x_{f_2}, b_1 and b_2 all are chosen to be greater than zero. Hence it is concluded that

$$\varphi_C \le \varphi_{NC} \tag{6.27}$$

Substitution of $q = 0$ in Eq. (6.25) results in no-cooperation. The desired decrement in the payoff function can be set with a judicious choice of the parameters used in the *z*-model.

6.7 Numerical Example: A 2-Pursuers and 1-Evader Cooperative Game

The error dynamics between two pursuers and one evader in the two-dimensional plane are

$$\dot{x}_1 = x_2; \dot{x}_2 = u_{P_1} - u_E \tag{6.28}$$

$$\dot{y}_1 = y_2; \dot{y}_2 = u_{P_2} - u_E \tag{6.29}$$

It is assumed that pursuer-1 (P_1) moves along the x-axis and similarly pursuer-2 (P_2) moves along the y-axis. Here, x_1 and x_2 are the relative position and velocity coordinates between P_1 and E, respectively. The relative position and velocity coordinates between P_2 and E are denoted by y_1 and y_2, respectively. Because the relative system description is linear, we consider a linear equation for the cooperation state z,

$$\dot{z} = az + c_1 x_1 + c_2 x_2 + c_3 y_1 + c_4 y_2 + b_1 u_{p_1} + b_2 u_{p_2} \tag{6.30}$$

Now we pose the dynamic game as follows,

$$\text{Min } J = \varphi + \frac{1}{2}\int_0^{t_f} (qz^2 + r_{P_1} u_{P_1}^2 + r_{P_2} u_{P_2}^2 - r_E u_E^2) dt \tag{6.31}$$

$$\text{where, } \varphi = \frac{1}{2}(s_{f_1} x_1^2 + s_{f_2} x_2^2 + s_{f_3} y_1^2 + s_{f_4} y_2^2 + s_c z^2) \tag{6.32}$$

The following parameter values are assumed:

$$\begin{aligned} &t_f = 3; x(0) = [-0.1\ 0\ 0.1\ 0]^T; q = r_{P_1} = r_{P_2} = \frac{1}{2}; r_E = 1; \\ &a = -1; c_1 = -2; c_2 = -0.1; c_3 = c_4 = -1; s_{f_1} = s_{f_3} = 2; s_{f_2} = s_{f_4} = 1; s_c = 0.5 \end{aligned} \tag{6.33}$$

These parameters result in the trajectories shown below in Fig. 6.7. The “outside” curves show the relative position coordinates for the no-cooperation case, whereas the “inside” curves show the relative position coordinates for the cooperation case. For this example, the payoff function, φ reflects the terminal miss, which decreases when it is evaluated with the cooperation. Figure 6.8 shows the evolution of the cooperation and it indicates

how the weights, r_{P_1} and r_{P_2} in the integrand given in Eq. (6.31) automatically changes the dynamical system with time to affect the final payoff.

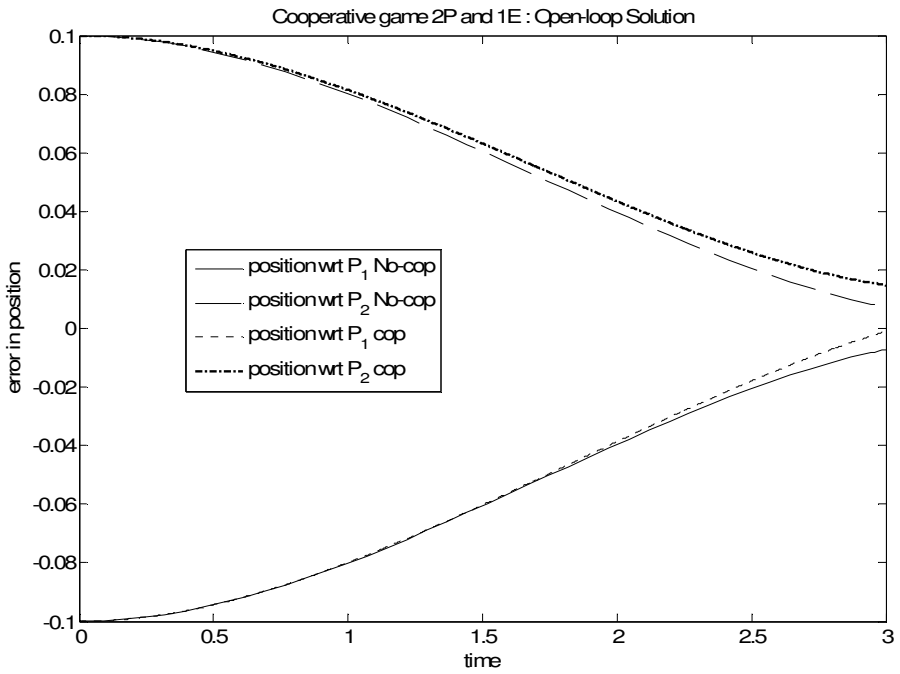


Fig. 6.7 Comparison between Cooperation and No Cooperation

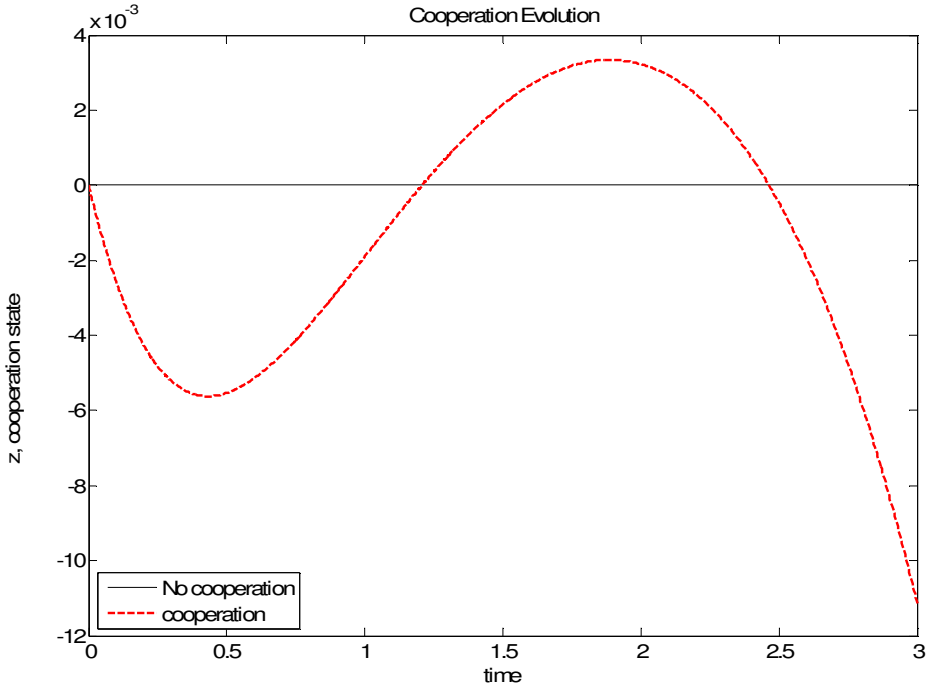


Fig. 6.8 Evolution of the z-model

CHAPTER VII

EXTENDED APPLICATIONS

In the previous chapters, SSM was applied to solve the continuous finite-horizon HJB/HJI equations with unconstrained control. This chapter provides further extensions and applications of SSM for a variety of other challenging optimal feedback control problems. In the first section of this chapter, feedback solutions of LQ problems with control bounds, posed as inequality constraints, are presented. The switching structure of the control and the obtained feedback solution are compared with their respective results obtained via analytical or the open-loop solutions.

Subsequent sections present an application to bilinear systems to systems with free parameters. A free final-time optimal control problem is considered to demonstrate the application of the SSM to optimize system parameters.

7.1 Optimal Feedback Control with Control Inequality Constraints

Solving optimal feedback control problems with control bounds is very challenging. Even for simple LQ problems with multiple control constraints, a general treatment to obtain the feedback solution is not possible because of the unknown control switching structure, non-smoothness of the value function, and the possibility of singular arcs in the solution. The following assumptions are made to simplify the problem:

1. Optimal solution exists in the given control domain $u \in U$ [admissible control region]
2. Switching structure is known a priori by using Pontryagin's principle[21].
3. Singular arcs[65] do not exist.

We begin with simple optimal control problems involving control bounds. Each problem is solved by using both the dynamic programming and calculus of variation approaches.

One-dimensional LQ Problem with Control Bounds

Consider the following one dimensional optimal control problem with no hard or soft constraint at the final time,

Min:

$$J = \int_{t_0}^{t_f} (-2x + 3u + u^2)dt$$

subject to (7.1)

$$\dot{x} = x + u; \qquad x(t_0) = 5; \quad x(t_f) = free$$

$$t_0 = 0; t_f = 2;$$

The control is bounded by,

$$0 \le u \le 2 \tag{7.2}$$

The above problem is chosen to demonstrate the steps for the application of dynamic programming along with Pontryagin's principle. This problem has an analytical solution which is derived below.

The Hamiltonian, H can be described as,

$$H = (u^2 + 3u - 2x) + \lambda(x + u) \tag{7.3}$$

where, λ is the costate vector.

Applying the necessary conditions for optimality and the Pontryagin's principle, the complete solution is

$$\lambda(t) = -2e^{-(t-2)} + 2 \tag{7.4}$$

and

$$u = \begin{Bmatrix} 2; & \lambda < -7 \\ -\frac{1}{2}(\lambda + 3) \equiv -\frac{1}{2}(5 - 2\mathrm{e}^{-(t-2)}); & \lambda \in [-7, -3] \\ 0; & \lambda > -3 \end{Bmatrix} \tag{7.5}$$

It can easily be shown that the control switches at the following instants:

$$t_{s_1} = 0.4959; t_{s_2} = 1.0837 \tag{7.6}$$

The optimal feedback solution to the above problem can be obtained from the HJB equation:

If the control is unbounded, the HJB equation is given as,

$$\frac{\partial J^*}{\partial t} = -\underbrace{H}_{\min u} = \frac{1}{4}\left\{\left(\frac{\partial J^*}{\partial x}\right)^2 + \frac{\partial J^*}{\partial x}(6 - 4x) + 8x + 9\right\}; J^*(x(t_f), t_f) = 0 \tag{7.7}$$

To solve Eq.(7.7), J^* is assumed to be a quadratic function of x with the time dependent gains s, p and c as,

$$J^* = \frac{1}{2}s(t)x^2 + p(t)x + c(t) \tag{7.8}$$

Substituting Eq. (7.8) into Eq. (7.7) yields the following gain differential equations,

$$\begin{aligned} \dot{s} &= \frac{1}{2}s^2 - 2s \text{ (Riccati Equation)} \\ \dot{p} &= \frac{1}{4}(2ps + 6s - 4p + 8) \\ \dot{c} &= \frac{1}{4}(p+3)^2 \end{aligned} \tag{7.9}$$

with the boundary conditions,

$$s(t_f) = 0; p(t_f) = 0; c(t_f) = 0\,. \tag{7.10}$$

Equation (7.9), with the use of Eq. (7.10), are solved analytically as,

$$\begin{aligned} s &= 0; \\ p &= 2(1 - e^{-(t-2)}); \\ c &= \left(17 - \frac{25}{4}t - 5e^{-(t-2)} + \frac{1}{2}e^{-2(t-2)}\right) \end{aligned} \tag{7.11}$$

Hence the feedback control, if the control is unbounded, can be obtained as shown below:

$$\begin{aligned} \lambda &= \frac{\partial J^*}{\partial x} = p = 2(1 - e^{-(t-2)}) \\ u &= -\frac{1}{2}(5 - 2e^{-(t-2)}) \end{aligned} \tag{7.12}$$

When control bounds are specified, the optimal control can either be within the bounded region or on the specified control boundaries. Hence, the gain differential equations on

the control boundaries have to be analyzed. There are two possibilities for the given control bounds, either u =0 or u =2. The respective HJB equations for the boundary values of control are given as,

i) $$\frac{\partial J^*}{\partial t} = x\left(2 - \frac{\partial J^*}{\partial x}\right) \text{ for } u = 0 \tag{7.13}$$

The series solution for the above HJB is,

$$\begin{aligned} &J^* = \frac{1}{2} s_1 x^2 + p_1 x \\ &\text{where} \\ &\dot{s}_1 = -2 s_1 \\ &\dot{p}_1 = 2 - p_1 \end{aligned} \tag{7.14}$$

ii) $$\frac{\partial J^*}{\partial t} = x\left(2 - \frac{\partial J^*}{\partial x}\right) - 2\frac{\partial J^*}{\partial x}(x+5) \text{ for } u = 2 \tag{7.15}$$

The series solution for the above case is,

$$\begin{aligned} &J^* = \frac{1}{2} s_2 x^2 + p_2 x + c_2 \\ &\text{where} \\ &\dot{s}_2 = -2 s_2 \\ &\dot{p}_2 = 2 - p_2 \\ &\dot{c}_2 = p_2 - 5 \end{aligned} \tag{7.16}$$

As it is known from open-loop solution that the continuity of the costate vector is maintained from initial time to the final time, continuity of the cost and its derivative with respect to x is assumed to evaluate the switch times. Hence,

$$\frac{\partial J^*}{\partial x} = 2 - 2e^{-(t_{s_1,s_2} - 2)} \tag{7.17}$$

From Eqs. (7.13)-(7.16), the gains at each switch time are,

$$p_1 = p_2 = -2e^{-(t_{s1,s2}-2)} \tag{7.18}$$

Figure 7.1 shows the optimal feedback trajectory whereas Fig.7.2 demonstrates optimal control histories for the unconstrained and constrained control cases, which are solved by using both the dynamic programming approach and the calculus of variations approach. It can easily be seen that both the approaches result in the same solution.

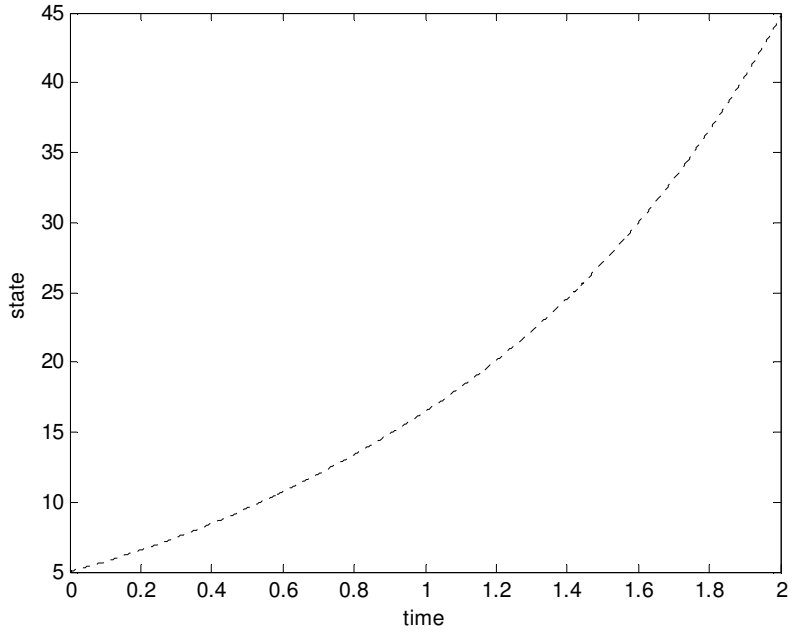


Fig. 7.1 Optimal Trajectory in the Presence of Control Constraints

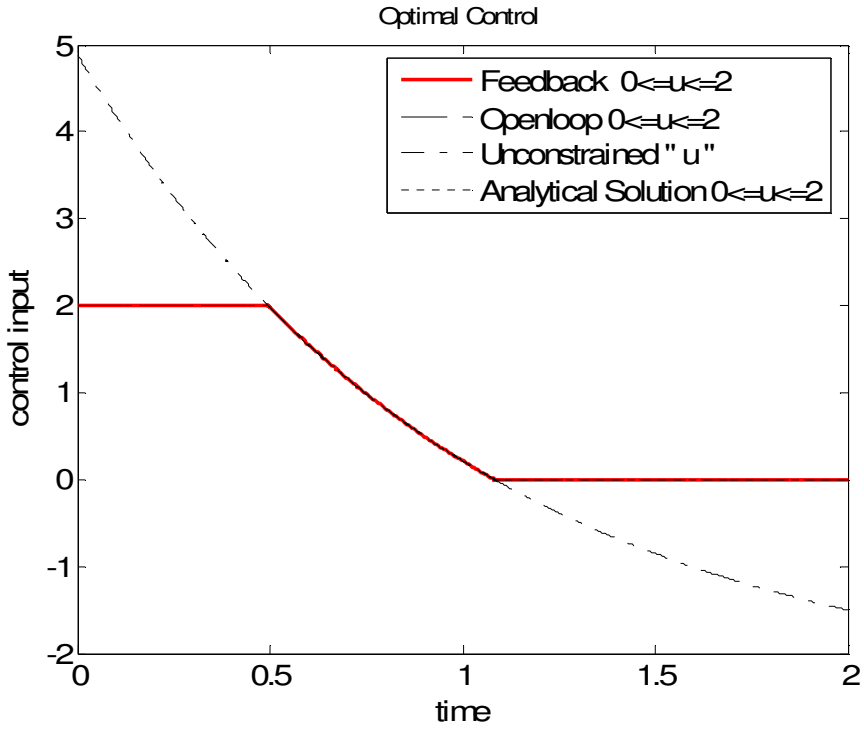


Fig. 7.2 Optimal Feedback Solution in the Presence of Control Constraints

LQ Problem with Bounded Control and a Terminal Constraint

The problem discussed above was that with a free final state. In this section, the same problem is treated with the addition of a terminal constraint. The problem with a point terminal constraint is posed as,

Min:

$$J = \int_0^2 (-2x + 3u + u^2)dt$$

subject to (7.19)

$$\dot{x} = x + u; \qquad x(0) = 5; \quad x(2) = 40$$

$$0 \le u \le 2; u \in U$$

For this problem, the respective HJB equations for unconstrained and constrained controls are the same as Eqs. (7.7),(7.13), and(7.15), only the control switches[65] can occur at different times and the number of switches may differ depending upon the values of the initial state and the terminal constraint. For some terminal constraints, since the control is bounded, even a feasible solution may not exist. Also there is a possibility that no switch or many switches are required to obtain the optimal solution.

In order to circumvent the issues mentioned above, an alternative waypoint scheme is implemented with an assumed control switching structure. An outer loop is used to minimize, the total cost with respect to the switching times. Then, with the known switching times, the inner loop can be processed by SSM and the known switching structure.

For the problem stated above, there could be three segments of the optimal trajectory, corresponding to admissible values of the control lying inside or on the constrained boundary. Therefore, the outer loop minimization problem can be stated as,

Minimize:

$$\begin{aligned} &\sum_i J_i^*(t_{s_1}, t_{s_2}) \\ &\text{subject to} \\ &0 \le \{t_{s_1}, t_{s_2}\} \le t_f \\ &u_{\min} \le \{u_{t_{s_1}}^*, u_{t_{s_2}}^*\} \le u_{\max} \end{aligned} \tag{7.20}$$

where, i is the number of segments in the problem and the superscript * denotes optimal values. Assuming a maximum of two control switches, additional optimization variables t_{s_1} and t_{s_2} are included in the outer loop. The Matlab® function “fmincon” is employed

to solve the minimization problem given by Eq. (7.20). The switch times are computed as,

$$t_{s_1} = 0; t_{s_2} = 0.56744 \tag{7.21}$$

Figures (7.3) and (7.4) show the open-loop and feedback solutions for the state and control, respectively. The solutions are indistinguishable from each other and the optimal control has a single switch which occurs at t_{s_2}, after which the control remains zero, till the end. For additional verification, the cost-to-go obtained by using the open-loop and feedback approaches are found to be the same:

$$J^*_{OL} = J^*_{FB} = -66.9502 \tag{7.22}$$

Note that, a general purpose program cannot be constructed for the outer loop optimization problem, since it requires the knowledge of the switching structure.

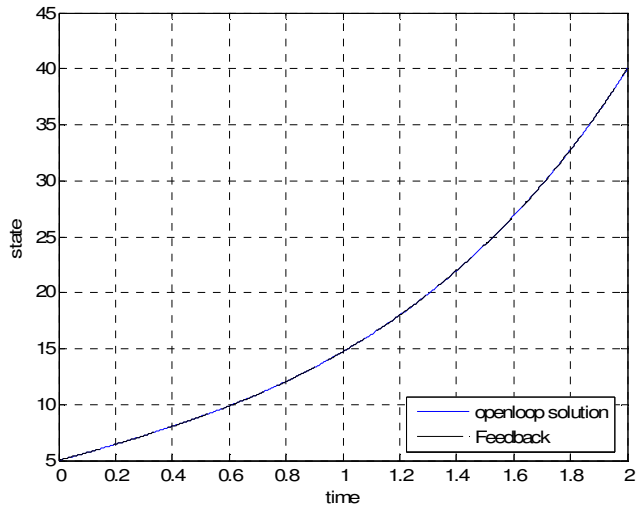


Fig. 7.3 Optimal State Trajectory for OCP with Control Bounds and a Terminal Constraint

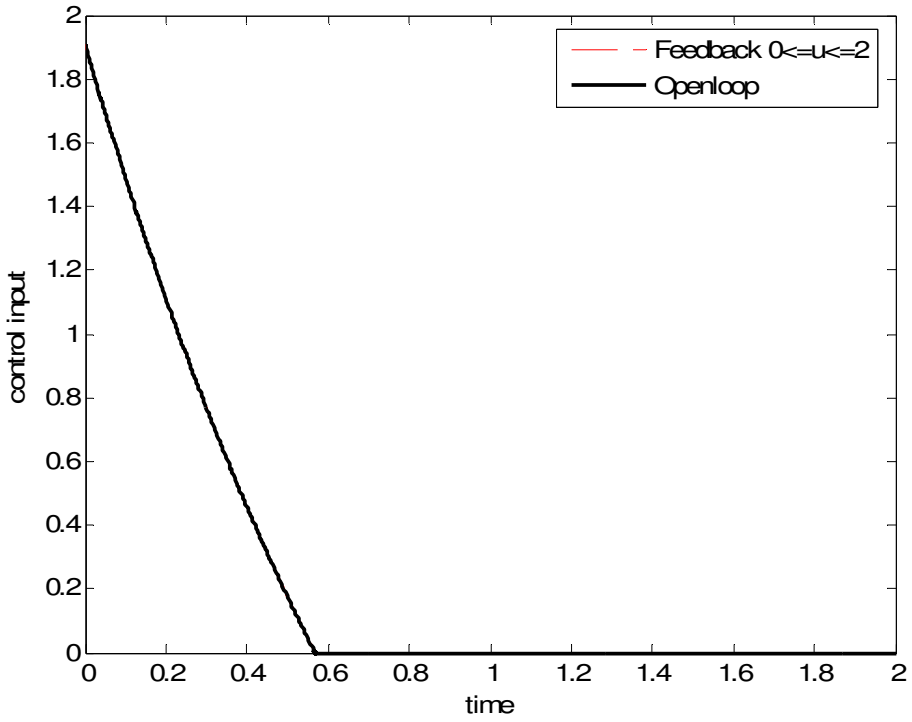


Fig. 7.4 Control Histories for the OCP with Control Bounds and a Terminal Constraint

A second example of an LQ problem with a terminal constraint is treated next:

Minimize:

$$
\begin{aligned}
&J = \int_0^5 u^2 dt \\
&\ddot{x} = u; \quad x(0) = 0.1; \quad \dot{x}(0) = 0.2 \\
&\psi = 2x(5) + \dot{x}(5) - 25 \\
&|u| \le 1.0
\end{aligned}
\tag{7.23}
$$

The open-loop solution for the initial costate vector, terminal constraint Lagrange multipliers and the cost, for the above problem are shown below in Table 7.1;

Table 7.1 Open-loop Solution with No Control Bounds and Control Bounds

	$\lambda(0)$	ν	J^*
No control bounds	[−0.2039; −1.1215]	−0.1020	1.1521
Control bounds	[−0.2090; −1.1494]	−0.1045	1.1540

Figure 7.5 and 7.6 show the feedback and open-loop solutions along with the solution of the OCP without control constraints. It can be observed that the feedback solution implemented by using SSM exactly follows the open-loop solution. Initially, the optimal control is at its maximum value and there is only one switch to satisfy the terminal constraint.

These problems are highly sensitive with respect to the switch times. With the same terminal constraint, the variation of the switch time as a function of the final time is studied in Fig. 7.7, which clearly demonstrates that the switch time is highly sensitive to changes in the final time. As the final time decreases, the state trajectory moves longer on the control saturation boundary and the switch time moves forward.

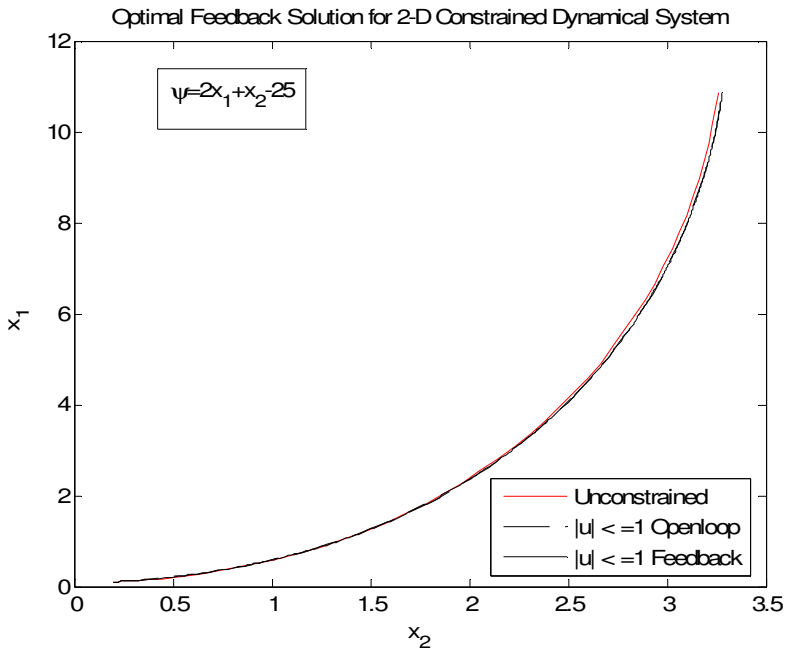


Fig. 7.5 Phase Plot for a 2-D LQ Problem with a Linear Constraint

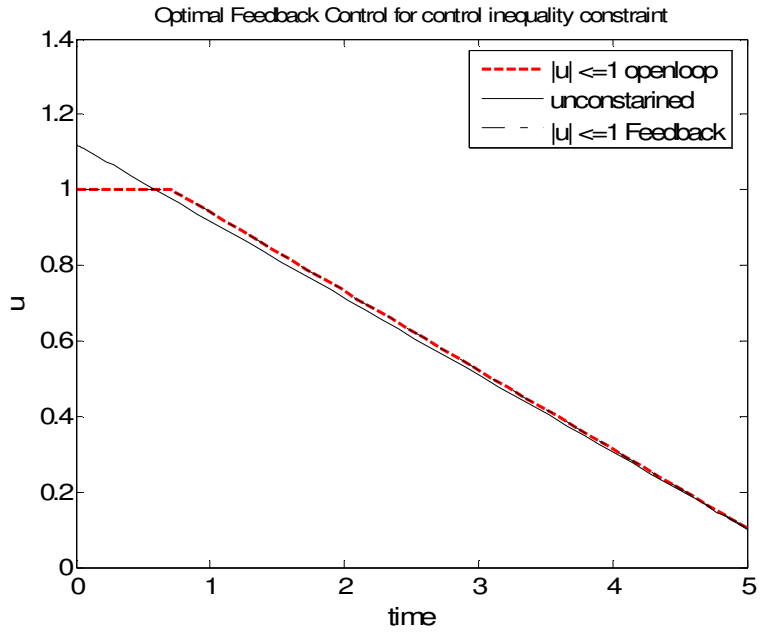


Fig. 7.6 Control Histories for a 2-D LQ Problem with a Linear Constraint

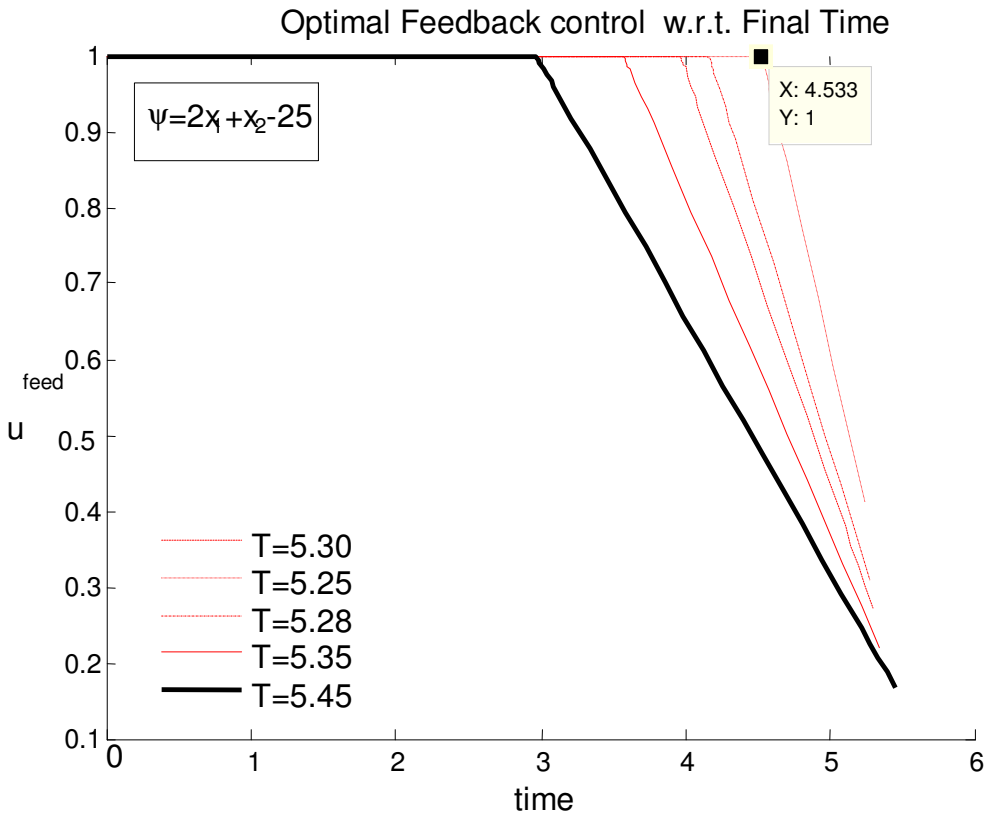


Fig. 7.7 Optimal Feedback Solution with Respect to the Final Time

7.2 Parameter Optimization

This section demonstrates the application of the SSM to solve the optimal control problems with parameters. A general optimal control problem with free parameter vector $p \in \Re^{p_1 \times 1}$ can be stated as,

Minimize:

$$J = \varphi + \int_{t_0}^{t_f} L(x,u,p)dt$$

subject to the system dynamics (7.24)

$$\dot{x} = f(x,u,p); x(t_0) = x_0$$

$$\psi(x(t_f)) = \psi_f$$

The parameter can be converted into a p_1 dimensional state by adding the state equation

$$\dot{p} = 0 \tag{7.25}$$

The Hamiltonian in the extended space can be given as,

$$H(x,u,p) = L + \lambda^T f + \mu^T [0] \tag{7.26}$$

The additional necessary conditions for optimality required to solve for the parameter are given as,

$$\dot{\mu} = -H_p \tag{7.27}$$

$$\Phi_p{}^T + \int_{t_0}^{t_f} H_p dt = 0 \tag{7.28}$$

where $\Phi = \varphi + v^T \psi$.

The following constraint is obtained by substituting Eq. (7.27) into Eq. (7.28);

$$\Phi_p\Big|^T_{t_f} + \int_{t_0}^{t_f} H_p dt = \Phi_p\Big|^T_{t_f} - \int_{t_0}^{t_f} \dot{\mu} dt = \Phi_p\Big|^T_{t_f} - \mu(t_f) + \mu(t_0) \tag{7.29}$$

As, *p* is unspecified at the initial time (or the final time), $\mu(t_0)$=0; which implies that,

$$\mu(t_f) = \Phi_p\Big|^T_{t_f} \tag{7.30}$$

The HJB equation for this problem is

$$\frac{\partial J^*}{\partial t} = -\underbrace{\min}_{u}\{H(x,u,p)\}; J^*(x(t_f)) = \varphi \quad \text{where } x(t_f) \text{ satisfies } \psi \tag{7.31}$$

As for the costate vector, λ μ can also be shown to satisfy the following condition:

$$\frac{\partial J^*(x,p,\nu,gains)}{\partial p} = \mu \tag{7.32}$$

Use of Eq. (7.30) in the above relation at the final time, results in

$$\frac{\partial J^*(x(t_f),\nu,p)}{\partial p} = \mu(t_f) = \Phi_p{}^T\Big|_{t_f} \tag{7.33}$$

In the HJB formalism, the solution procedure for the parameter optimization problem can be described in the following steps.

1. Consider the SSM in the extended state space, $x \in \Re^{(n+p_1)\times 1} \equiv \{x, p\}$ as $J^* = J^*(x, p, \nu, gains(t))$. The boundary conditions for the additional gains associated with monomials of the parameter vector, p can be obtained by using Eq. (7.33).
2. Integrate the gain differential equation backward in time and store the gains, Determine p and ν from the following equations:

$$\frac{\partial J^*(x(t_0),\nu,p)}{\partial v} = 0; \qquad \frac{\partial J^*(x(t_0),\nu,p)}{\partial p} = \varphi_p + \psi_p^T \nu \tag{7.34}$$

Note that the system of equations, Eq. (7.34) can be nonlinear in p and ν.

3. Use the obtained value of the parameter vector from the above step while integrating the closed-loop system forward in time. (Terminal Lagrange multiplier and parameter vector can be updated for nonlinear problems).

7.3 Feedback Solution of Free Final Time Optimal Control Problems

The free final time problem can be converted into a fixed final time problem by introducing the nondimensional coordinates for time:

$$\tau = \frac{t - t_0}{t_f - t_0} \tag{7.35}$$

where $\tau(t_0) = 0; \tau(t_f) = 1,$

Then, the optimal control problem can also be restated in terms of the parameter t_f. To demonstrate the applicability of SSM to the parameter optimization problem, the following LQ problem with a point terminal constraint is considered,

$$\begin{aligned}
&Min: \\
&J = \int_{t_0}^{t_f} \left(2x + \frac{1}{2}u^2 \right) dt \\
&\text{subject to} \\
&\dot{x} = u \\
&\psi(t_f) \equiv x(t_f) = \psi_f = 4 \\
&t_0 = 1, x(t_0) = 4 \\
&t_f \text{ is free, } x \in \Re^1; u \in \Re^1; \frac{d}{dt} \equiv (\dot{}\,)
\end{aligned} \tag{7.36}$$

With the parameter $p = t_f$, by using Eq.(7.35), the problem (7.36) is converted into a new OCP, as follows:

Minimize

$$J = \int_0^1 (p-1)\left(2x + \frac{1}{2}u^2\right) d\tau$$

subject to

$$\begin{aligned} &x' = -u + pu; \\ &x(0) = 4; \\ &\psi(1) \equiv x(1) = \psi_f = 4; \\ &\mathbf{x} \in \Re^2 \equiv [x, p]; u \in \Re^1; \frac{d}{d\tau} \equiv (') \end{aligned} \tag{7.37}$$

Solution Procedure

In the modified form, the free-final time LQ problem appears as a fixed-final time bilinear system with a free parameter p. The HJB equation for the above OCP is,

$$\frac{\partial J^*}{\partial t} + (p-1)\left[2x - \frac{1}{2}\left(\frac{\partial J^*}{\partial x}\right)^2\right] = 0 \tag{7.38}$$

where J^* is the value function.

A power series expansion of J^* in x, p and ν is considered up to sixth order in p (because it appears bilinearly) and second order in x and ν. Substituting the assumed J^* into the HJB equation (7.38) yields 63 gain differential equations and the known transversality conditions for optimality [3,65] return the boundary conditions to solve each of them numerically. All gains can be stored by integrating them backwards in nondimensional time.

For obtaining the initial value of ν and the optimal value of p, Eq. (7.34) is solved at $\tau = 0$ by using the Matlab® routine "fsolve", which returns

$$p = t_f = 5.00000000000023 \tag{7.39}$$

$$\nu(0) = -4.00000000000045 \tag{7.40}$$

The obtained solution for the parameter matches with the analytical optimal solution which is $t_f = 5$. The analytical solution for the state is

$$x^*(t) = t^2 - 6t + 9 \,; t_f = 5 \tag{7.41}$$

Fig. 7.8 shows the plots of the optimal trajectories obtained by using the analytical solution and the SSM.

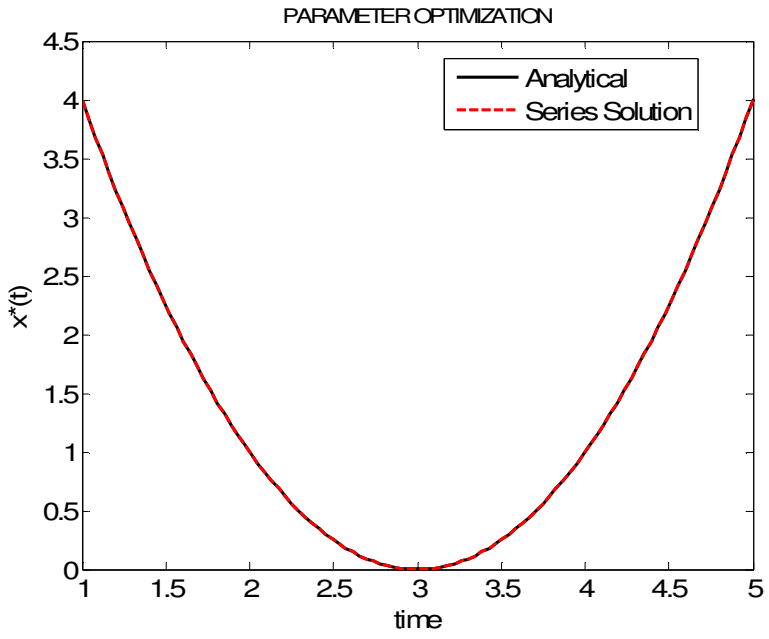


Fig. 7.8 Feedback Solution of Free Final Time Optimal Control Problem

CHAPTER VIII

CONCLUSIONS

This dissertation has presented a series solution framework for solving the HJB/HJI equations associated with finite-horizon optimal feedback control problems for nonlinear systems with nonlinear terminal constraints, H-infinity control and cooperative pursuit-evasion games. Two methods developed in this dissertation are the Series Solution Method (SSM) and the Galerkin Approximation Technique (GAT).

SSM is based upon the Taylor' series expansion of the value function in terms of the states and the constraint Lagrange multipliers. The HJB/HJI partial differential equation is converted into a system of ordinary differential equations (ODEs) for the coefficients of the series expansion by collecting coefficients of like powered terms. SSM provides a higher-order approximation to the solution of the HJB/HJI equation with high local accuracy. GAT utilizes orthogonal polynomial basis functions for the expansion of the value function and the required ODEs are obtained via the projection approach and, it achieves a uniform level of error over the computational domain.

SSM requires a polynomial form of the Hamiltonian and is best suited when the dynamical system has linear terms. On the other hand, GAT is applicable to systems with more general, smooth nonlinear structures. The extension of SSM to higher spatial

dimensions is relatively easier as compared to GAT, which requires the evaluation of multi-dimensional spatial integrals.

A waypoint scheme has been derived to broaden the applicability of SSM over large domains of time. This approach, equivalent to a feedback implementation of the multiple-shooting method, is capable of reducing the gain storage requirements, especially for autonomous systems. The waypoint scheme has been shown to produce significantly less midcourse error than that by the direct SSM.

SSM has been applied to the problems of nonlinear differential and cooperative games with nonlinear terminal constraints. Several numerical examples are presented to show the efficacy of the method for obtaining Pareto optimal feedback solutions. An extension of the SSM to the solution of cooperative game (cooperation among the pursuers) problems is accomplished via the development of a new z-model. It is shown that the developed algorithm provides a structured approach to affect the capture of a powerful evader.

Finally, SSM has been extended to handle optimization problems involving control inequality constraints and free parameters.

REFERENCES

1. Dreyfus, S., *Dynamic Programming and the Calculus of Variations*, Academic Press, New York, 1965.
2. Bellman, R.E., and Dreyfus, S. E., *Applied Dynamic Programming*, 1962, Princeton University Press, Princeton, NJ.
3. Bryson, A. E., *Dynamic Optimization*, Addison Wesley Longman Inc., 1999.
4. Bellman, R.E., *Dynamic Programming*, 1957, Princeton University Press, Princeton, NJ.
5. Basar, T. S. and Bernhard, P., "Differential Games and Applications", *Lecture Notes in Control and Information Sciences*, Vol. 119, Springer-Verlag, 1989.
6. Isaacs, R., *Differential Games: A Mathematical Theory with Applications to Warfare and Pursuit, Control and Optimization*, John Wiley and Sons, Inc, New York, 1965.
7. Isaacs, R., "Differential Games: Their Scope, Nature, and Scope," *Journal of Optimization Theory and Applications*, Vol. 3, No.5, 1969, pp. 283-295.
8. Basar, T. S. and Olsder, G. J., *Differential Noncooperative Game Theory*, SIAM, Philadelphia, PA, 1999.
9. Engwerda, J. C., *LQ Dynamic Optimization and Differential Games*, Wiley, Chichester, UK, 2005.
10. Broek, W. A. van den, "Uncertainty in Differential Games," Ph.D. Thesis, Center Dissertation Series, Tilburg University, The Netherlands, 2001.
11. Al'brekht, E. G., "On the Optimal Stabilization of Nonlinear Systems", *Journal of Applied Mathematics and Mechanics*, Vol.25, 1962, pp. 1254-1266.
12. Lukes, D. L. "Optimal Regulation of Nonlinear Dynamical Systems", *SIAM Journal of Control*, Vol.7, No.1, 1969, pp.75-100.

13. Dabbous, T. E. and Ahmed, N. U., "Nonlinear Optimal Feedback Regulation of Satellite Angular Momenta," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-18, No. 1, Jan. 1982, pp. 2-10.
14. Carrington, C. K. and Junkins, J. L., "Optimal Nonlinear Feedback Control for Spacecraft Attitude Maneuvers," *Journal of Guidance, Control, and Dynamics*, Vol. 9,No. 1, 1986, pp. 99-107.
15. Yoshida, T. and Laparo, K. A., "Quadratic Regulatory Theory for Analytic Nonlinear Systems with Additive Controls", *Automatica*, Vol.25, No.4, 1989, pp. 425-435.
16. Garrard, W. L., McClamroch, N. H. and Clark, L. G., An Approach to Suboptimal Feedback Control of Nonlinear Systems", *International Journal of Control*, Vol.5, No.5, 1967, pp. 425-435.
17. Cloutier, J. R., State-Dependent Riccati Equation Techniques, *Proceedings of the American Control Conference*, Albuquerque, NM, June 1997.
18. Tewari, A., Optimal Nonlinear Spacecraft Attitude Control through Hamilton-Jacobi Formulation, *Journal of the Astronautical Sciences*, Vol. 50, No. 1, 2002, pp. 99-112.
19. Sharma, R. and Tewari, A., Optimal Nonlinear Tracking of Spacecraft Attitude Maneuvers, *IEEE Transactions on Control Systems Technology*, Vol. 12, No. 5, Sep. 2004, pp. 677-682.
20. Navasca, C. L., and Krener, A. J., "Solutions of Hamilton Jacobi Bellman Equations", *Proceedings of the 39th IEEE Conference on Decision and Control*, Sydney, Australia, Dec.2000, Vol.1, pp. 570-574.
21. Pontryagin, L.S., Boltyanskii, V. G., Gamkrelidge, R. V. and Mischenko, E. F., *The Mathematical Theory of Optimal Processes*, John Wiley and Sons, New York, 1962.
22. Cerven, W. T. and Bullo, F., "Constructive Controllability Algorithms for Motion Planning and Optimization," *IEEE Transactions on Automatic Control*, Vol. 48, No.4, April 2003, pp. 575-588.

23. Bullo, F. and Cerven, W. T., "On Trajectory Optimization for Polynomial Systems via Series Expansions," *Proceedings of the 39th IEEE Conference on Decision and Control*, Sydney, Australia, December 2000.
24. Lyshevski, S. E., "Optimal Control of Nonlinear Continuous-time Systems: Design of Bounded Controllers via Generalized Nonquadratic Functionals," *Proceedings of the American Control Conference*, Philadelphia, PA, June 1998, pp. 205-209.
25. Cimen, T. and Banks, S. P., "Global Optimal Feedback Control for General Nonlinear Systems with Nonquadratic Performance Criteria," *Systems & Control Letters*, Vol. 53, No. 5, 2004, 327-424.
26. Xin, M. and Balakrishnan, S. N., "Nonlinear Missile Autopilot Design with $\theta - D$ Technique," *Journal of Guidance, Control, and Dynamics,* Vol. 27, No. 3, 2004, pp. 406-417.
27. Beard, R., Saridis, G., and Wen, J., "Galerkin Approximation of the Generalized Hamilton-Jacobi-Bellman Equation," *Automatica*, Vol. 33, No. 12, 1997, pp. 2159-2177.
28. Beard, R., Saridis, G., and Wen, J., "Improving the Performance of Stabilizing Controls for Nonlinear Systems," *IEEE Control Systems Magazine*, Vol. 16, No. 5, 1996, pp. 27-35.
29. Beard, R. and McLain, T., "Successive Galerkin Approximation Techniques for Nonlinear Optimal and Robust Control," *International Journal of Control*, Vol. 71, No. 5, 1998, pp. 717-748.
30. Park, C. and Tsiotras, P., "Sub-optimal Feedback Control Using a Successive Wavelet-Galerkin Algorithm," *Proceedings of the American Control Conference*, Dayton, OH, June 2003, pp. 1926-1931.
31. Park, C. and Tsiotras, P., "Approximations to Optimal Feedback Control Using a Successive Wavelet Collocation Algorithm," *Proceedings of the American Control Conference*, Dayton, OH, June 2003, pp. 1950-1955.

32. Richardson, S. and Wang, S., "Numerical Solution of Hamilton-Jacobi-Bellman Equations by an Exponentially Fitted Finite Volume Method 1," *Optimization*, Vol. 55, No. 1&2, Feb 2006, pp. 121-140.
33. Jin, S. and Osher, S., "A Level Set Method for the Computation of Multivalued Solutions to Quasi-linear Hyperbolic PDEs and Hamilton-Jacobi Equations," *Comm. Math. Sci*. Vol. 1, No. 3, 2003, pp. 575-591.
34. Huang, Yun and Lu Wei-min, "Nonlinear Optimal Control: Alternatives to Hamilton-Jacobi Equation," *Proceedings of the 35th Conference on Decision and Control*, Kobe, Japan, Dec 1996, pp. 3942-3947.
35. Bayen, A. M. and Tomlin, C. J., "A Construction Procedure Using Characteristics for Viscosity Solutions of the Hamilton-Jacobi Equation," *Proceedings of the 48th IEEE Conference on Decision and Control*, Orlando, Florida, USA, Dec. 2001, pp. 1657-1662.
36. Bardi,M and Capuzzo-Dolcetta, I., *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*, Birkhauser Publications, Boston, 1997.
37. Guibout, V. M. and Scheeres, D. J., "Solving Relative Two-point Boundary Value Problems: Spacecraft Formation Flight Transfer Applications," *Journal of Guidance, Control, and Dynamics*, Vol. 27, No. 4, 2004, pp. 693-704.
38. Park, C. and Scheeres, D. J., "Solutions of Optimal Feedback Control Problems with General Boundary Conditions Using Hamiltonian Dynamics and Generating Functions," *Proceedings of the American Control Conference*, Boston, Massachusetts, June 2004, Paper WeM02.1.
39. Park, C. and Scheeres, D. J., "Extended Applications of Generating Functions to Optimal Feedback Control Problems," *Proceedings of the American Control Conference*, Dayton, OH, Vol. 2, June 2005, pp. 852-857.

40. Huang, J., and Lin, C.F., "Numerical Approach to Computing Nonlinear H-infinity Control Laws" , *Journal of Guidance, Control and Dynamics*, Vol.18, No.5, 1995, pp. 989-994.
41. Huang, J., and Lin, C. F., "Flight Control Design by Nonlinear Servomechanism Theory", *Proceedings of the American Control Conference*, San Francisco, CA, Vol. 1,1993, pp. 410-414.
42. Tsiotras, P., Corless, M and Rotea, M, "An L2 Disturbance Attenuation Solution to the Nonlinear Benchmark Problem", *IMA Journal of Math., Control and Information,* Vol.8, 1998, pp. 330-330.
43. Wise, K. A., and Sedwick, L. J., "Successive Approximation Solution of the HJI Equation", *Proceedings of the 33rd Conference on Decision and Control*, Lake Buena Vista, FL, Dec. 1994, pp. 1387-1391.
44. Murad, Abu-Khalaf and Lewis, F. L., "Nearly Optimal Control Laws for Nonlinear Systems with Saturating Actuators Using a Neural Network HJB Approach", *Automatica*, Issue 5, 2005, pp.779-791.
45. Murad Abu-Khalaf, Huang, J., and Lewis, F. L., Nonlinear H_2 / H_∞ Constrained Feedback Control, 2006, Springer-Verlag Limited, Germany.
46. Murad, Abu-Khalaf, Lewis, F. L. and Huang, J., "Hamilton-Jacobi-Isaacs Formulation for Constrained Input Nonlinear Systems", *Proceedings of the 43rd IEEE Conference on Decision and Control*, Atlantis, Paradise Island, Bahamas, Dec. 14-17, 2004
47. Cheng, T., Lewis, F. L., and Murad, A., "A Neural Network Solution for Fixed-Final Time Optimal Control of Nonlinear Systems," *Automatica*, Vol. 43, 2007, pp. 482-490.
48. Soravia, P., "H-infinity Control of Nonlinear Systems: Differential Games and Viscosity Solutions", *Siam Journal of Control and Optimization*. Vol. 34, No. 3, 1996, pp. 1071-1097.

49. van der Schaft, A. J., "A State-Space Approach to Nonlinear H_∞ Control", *Systems and Control Letters*, Vol. 16, 1991, pp. 1-8.
50. Sakamoto, N. and van der Schaft, A. J., " An Approximation Method for Stabilizing Solution of the Hamilton-Jacobi Equation for Integrable Systems Using Hamiltonian Perturbation Theory," *Proceedings of the 45th IEEE Conference on Decision and Control*, San Diego, CA, USA, Dec. 13-15, 2006, pp. 5857-5862.
51. Subrahmanyam, M. B., *Finite Horizon H_∞ and Related Control Problems*, Systems and Control: Foundations and Applications, 1995, Birkhauser, Boston.
52. Starr, A. W. and Ho, Y.-C., "Nonzero-Sum Differential Games," *Journal of Optimization Theory and Applications*, Vol. 3, No.3, 1969, pp. 185-206.
53. Doyle, C. J., Glover, K., Khargonekar, P.P, and Francis, B. A., "State-Space Solutions to Standard H_2 and H_∞ Control Problems", *IEEE Transactions on Automatic Control*, Vol. 34, No.8, August 1989, pp. 831-847.
54. Lancaster, P. and Rodman, L., *Algebraic Riccati Equations*, Oxford, Oxford University Press, 1995.
55. Isidori, A., and Astolfi, A., "Disturbance Attenuation and H_∞-Control Via Measurement Feedback in Nonlinear Systems", *IEEE Transactions on Automatic Control*, Vol. AC-37, No. 9, 1992, pp. 1283-1293
56. Basar, T. S. and Bernhard, P., "Differential Games and Applications", *Lecture Notes in Control and Information Sciences*, Vol. 119, Springer-Verlag, 1989.
57. Shima, T., and Shinar, J., "Time Varying Linear Pursuit-Evasion Game Models with Bounded Controls," *Journal of Guidance, Control, and Dynamics*, Vol. 25, No. 3, May-June 2002, pp. 425–431.
58. Horie, K. and Conway, B. A., "Optimal Fighter Pursuit-Evasion Maneuvers Found via Two-sided Optimization," *Journal of Guidance, Control, and Dynamics*, Vol. 29, No. 1, January-Feb 2006, pp. 105–112.

59. Breitner, M. H., Pesch, H. J., and Grimm, W., "Complex Differential Games of Pursuit–Evasion Type with State Constraints, Part 1: Necessary Conditions for Open-Loop Strategies," *Journal of Optimization Theory and Applications*, Vol. 78, No. 3, 1993, pp. 419–441.
60. Starr, A. W. and Ho, Y.-C., "Further Properties of Nonzero-Sum Differential Games," *Journal of Optimization Theory and Applications*, Vol. 3, No.4, 1969, pp. 207-219.
61. Nash, J., "Two-Person Cooperative Game," *Econometrica*, Vol. 21, 1953, pp. 128-140
62. Nash, J., "The Bargaining Problem," *Econometrica*, Vol. 18, 1950, pp. 155-162
63. Thomson, W., "Cooperative Models of Bargaining," *In Aumann and Hart (Eds.), Handbook of Game Theory*, Springer-Verlag Limited, Vol. 2. 1994, pp. 1238-1277.
64. Bryson, A. E. and Ho, Y.-C., *Applied Optimal Control*, Hemisphere Publishing Co., Washington, D.C., 1975.
65. Lewis, F.L. and Syrmos, V.L., *Optimal Control*, John Wiley & Sons, Inc., New York, 1995.
66. Turner, J. D., "Generalized Gradient Search and Newton's Method for Multilinear Algebra Root Solving and Optimization Applications," Invited Paper No. AAS 03-261, *Proceedings of John L. Junkins Astrodynamics Symposium, George Bush Conference Center*, College Station ,Texas, May 23-24, 2003.
67. Huang,J., and Lin, C.F., "Numerical Approach to Computing Nonlinear H-infinity Control Laws," *Journal of Guidance, Control and Dynamics*, Vol.18, No.5, pp. 989-994, 1995
68. Betts, J. T., "A Survey of Numerical Methods for Trajectory Optimization," *Journal of Guidance, Control, and Dynamics*, Vol. 21, 1998, pp. 183-207.

69. Stipanovic, D., Inhalan, Teo, R. and Tomlin, C.J., "Decentralized Overlapping Control of a Formation of Unmanned Aerial Vehicles," *Automatica*, Vol. 40, No. 8, August 2004, and pp. 1285-1296
70. Tomlin, C.J. and Sastry, S., "Conflict Resolution for Air Traffic Management: A Study in Multi-agent Hybrid Systems," *IEEE Transactions on Automatic Control*, Vol.43, No.4, April 1998, pp. 509-521.
71. Raivio, T. and Ehtamo, H., "Visual Aircraft Identification as a Pursuit-Evasion Game," *Journal of Guidance, Control and Dynamics*, Vol. 23, No. 4, July-August 2000, pp. 701-708
72. Aarle, B. van, Bovenberg, L. and Raith, M., "Monetary and Fiscal Policy Iteration and Government Debt Stabilization," *Journal of Economics*, Vol. 62, pp. 111-140.
73. Vadali, S. R. and Sharma, R., "Optimal Finite-time Feedback Controllers for Nonlinear Systems with Terminal Constraints," *Journal of Guidance, Control, and Dynamics*, Vol. 29, No. 4, July-August 2006, pp. 921–928.
74. Matlab®, http://www.mathworks.com/access/helpdesk/help/helpdesk.html
75. Junkins, J.L. and Turner, J.D., *Optimal Spacecraft Rotational Maneuvers*, Elsevier Science Publishing Co. New York, 1986.

APPENDIX A

A SYMBOLIC PROCEDURE FOR THE APPLICATION OF THE SERIES SOLUTION METHOD

The effect of the "curse of dimensionality" stated by Bellman can be reduced if the dynamic programming approach to solve the HJB equation is programmed elegantly. Representation of gains in a structured format, taking advantage of symmetries of the tensors, can minimize storage requirements. This appendix presents the development of a symbolic code for the application of the SSM.

A.1 Continuous Finite-time Optimal Control Problem with Terminal Constraints

A somewhat general continuous optimal control problem involving polynomial nonlinearities is given as,

Minimize:

$$J = \phi(x(t_f) + \int_{t_0}^{t_f} L(x,u)dt \tag{A.1}$$

Subject to

$$\dot{x}_i = f_i(x,u)\,; \tag{A.2}$$

$$\psi_i = g_i(x) - (\psi_f)_i \tag{A.3}$$

where,

$$L(x,u)=\left(C_J+q_{1_i}x_i+\frac{1}{2}q_{2_{ij}}x_ix_j+\frac{1}{3}q_{3_{ijk}}x_ix_jx_k+\frac{1}{4}q_{4_{ijkl}}x_ix_jx_kx_l+\frac{1}{2}r_{ij}u_iu_j\right);r_{ij}>0 \tag{A.4}$$

and

$$f_i(x,u)=\left(a_{1_{ij}}x_j+a_{2_{ijk}}x_jx_k+a_{3_{ijkl}}x_jx_kx_l+a_{4_{ijklm}}x_jx_kx_lx_m+C_i+b_{ij}u_j\right) \tag{A.5}$$

where $C_i, q_{1_i}, a_{1_{ij}}$ ..etc. are known coefficients.

As can be seen, the state equations and the performance index contain nonlinearities up to the 4th order in the states. The controls appear linearly in the state equations and quadratically in the performance index.

The Hamiltonian for the above OCP can be described as,

$$\begin{aligned}H=&\left(C_J+q_{1_i}x_i+\frac{1}{2}q_{2_{ij}}x_ix_j+\frac{1}{3}q_{3_{ij}}x_ix_jx_k+\frac{1}{4}q_{4_{ij}}x_ix_jx_kx_l+\frac{1}{2}r_{ij}u_iu_j\right)+\\&+\lambda_i\left(a_{1_{ij}}x_j+a_{2_{ijk}}x_jx_k+a_{3_{ijkl}}x_jx_kx_l+a_{4_{ijklm}}x_jx_kx_lx_m+C_i+b_{ij}u_j\right)\end{aligned} \tag{A.6}$$

Applying the necessary condition for optimality results in,

$$\frac{\partial H}{\partial u_i}=0 \tag{A.7}$$

$$\Rightarrow r_{ij}u_i+\lambda_jb_{ji}=0 \tag{A.8}$$

$$\Rightarrow u_i=-r_{ij}^{-1}b_{kj}\lambda_k;u_j=-r_{iq}^{-1}b_{mq}\lambda_m \tag{A.9}$$

The HJB Equation for the above OCP is shown below:

$$-\frac{\partial J^*}{\partial t}=\underbrace{\min}_{u_i}[H(x_i,u_i,\lambda_i)], \tag{A.10}$$

where the costate vector is,

$$\lambda_i = \frac{\partial J^*}{\partial x_i} \tag{A.11}$$

A.2 Development of the Series Solution Method by Using Indicial Notation

The following development utilizes indicial notation. The cost-to-go is expanded as shown below:

$$\begin{aligned} J^* = {} & p_0 + \left(p_i x_i + p_{1_i} \nu_i \right) + \left(\frac{1}{2} s_{ij} x_i x_j + p_{5_{ij}} x_i \nu_j + \frac{1}{2} p_{2_{ij}} \nu_i \nu_j \right) + \\ & + \left(\frac{1}{3} t_{ijk} x_i x_j x_k + p_{6_{ijk}} \nu_i \nu_j x_k + p_{8_{ijk}} \nu_i x_j x_k + \frac{1}{3} p_{3_{ijk}} \nu_i \nu_j \nu_k \right) + \\ & + \left(\frac{1}{4} g_{ijkl} x_i x_j x_k x_l + p_{7_{ijkl}} \nu_i \nu_j \nu_k x_l + p_{9_{ijkl}} \nu_i \nu_j x_k x_l + p_{10_{ijkl}} \nu_i x_j x_k x_l + \frac{1}{4} p_{4_{ijkl}} \nu_i \nu_j \nu_k \nu_l \right) \end{aligned} \tag{A.12}$$

A direct substitution of Eq. (A.12) into Eq.(A.11) results in the following:

$$\begin{aligned} \lambda_i = {} & p_i + s_{ij} x_j + t_{ijk} x_j x_k + g_{ijkl} x_j x_k x_l + p_{5_{ij}} \nu_j + p_{6_{jki}} \nu_j \nu_k + p_{7_{jkli}} \nu_j \nu_k \nu_l + 2 p_{8_{jik}} \nu_j x_k + \\ & + 2 p_{9_{jkil}} \nu_j \nu_k x_l + 3 p_{10_{jikl}} \nu_j x_k x_l \end{aligned} \tag{A.13}$$

Another key equation to evaluate the terminal Lagrange multiplier is,

$$\begin{aligned} \frac{\partial J^*}{\partial \nu_i} \equiv 0 = {} & p_{1_i} + \left(p_{2_{ij}} \nu_j + p_{5_{ji}} x_j \right) + \left(p_{3_{ijk}} \nu_j \nu_k + 2 p_{6_{ijk}} \nu_j x_k + p_{8_{ijk}} x_j x_k \right) + \\ & + \left(p_{4_{ijkl}} \nu_j \nu_k \nu_l + 3 p_{7_{ijkl}} \nu_j \nu_k x_l + 2 p_{9_{ijkl}} \nu_j x_k x_l + p_{10_{ijkl}} x_j x_k x_l \right) \end{aligned} \tag{A.14}$$

The Hamiltonian can be obtained by substituting Eq.(A.9) into Eq. (A.6):

$$\begin{aligned} H = {} & \left(C_J + q_{1_i} x_i + \frac{1}{2} q_{2_{ij}} x_i x_j + \frac{1}{3} q_{3_{ij}} x_i x_j x_k + \frac{1}{4} q_{4_{ij}} x_i x_j x_k x_l \right) \\ & - \left(\frac{1}{2} r_{ij} r_{ip}^{-1} r_{jq}^{-1} b_{kp} b_{mq} \frac{\partial J^*}{\partial x_k} \frac{\partial J^*}{\partial x_m} \right) + \left(\frac{\partial J^*}{\partial x_i} \right) \left(a_{1_{ij}} x_j + a_{2_{ijk}} x_j x_k + a_{3_{ijkl}} x_j x_k x_l + a_{4_{ijklm}} x_j x_k x_l x_m + C_i \right) \end{aligned}$$

(A.15)

The symmetry properties of the gain tensors used for the expansion of J^* are discussed in the next section.

A.3 Properties of the Gain Tensors

It is well known that SSM utilizes a minimal number of all possible monomials of a given order. The symmetry properties of the tensors are listed below:

$$(i) s_{ij} = s_{ji}$$
$$(ii) t_{ijk} = t_{jik} = t_{jki}$$
$$(iii) g_{ijkl} = g_{jikl} = g_{jkil} = g_{jkli}$$

$$(iv) p_{2_{ij}} = p_{2_{ji}}$$
$$(v) p_{3_{ijk}} = p_{3_{jik}} = p_{3_{jki}}$$
$$(vi) p_{4_{ijkl}} = p_{4_{jikl}} = p_{4_{jkil}} = p_{4_{jkli}}$$
$$(vii) p_{6_{ijk}} = p_{6_{jik}}$$

$$(viii) p_{7_{ijkl}} = p_{7_{jikl}} = p_{7_{jkil}}$$
$$(ix) p_{8_{ijk}} = p_{8_{ikj}}$$
$$(x) p_{9_{ijkl}} = p_{9_{jikl}}; p_{9_{ijkl}} = p_{9_{ijlk}}$$
$$(xi) p_{10_{ijkl}} = p_{10_{ikjl}} = p_{10_{iklj}}$$

(A.16)

For the n-dimensional system dynamic model, the total number of gain terms associated with the respective higher-order tensor can be calculated by using these general formulae:

$$s_{ij} = \frac{n(n+1)}{2}$$
$$s_{ijk} = \frac{n(n+1)(n+2)}{3!}$$
$$t_{ijkl} = \frac{n(n+1)(n+2)(n+3)}{4!} \tag{A.17}$$

For example, if the dimension of the dynamical system is six and the terminal constraint also belongs to $\Re^{6\times 1}$, then the number of terms in the respective matrices and tensors can be calculated as follows:

i) Gain Matrices (with symmetry):

$$s = 21, p_2 = 21 \tag{A.18}$$

ii) Gain Vectors and Matrices (without symmetry)

$$p = 6, p_1 = 6, p_5 = 36, p_0 = 1 \tag{A.19}$$

iii) Other Higher order tensors (with full and partial symmetry)

$$t_{ijk} = 56, p_3 = 56, p_6 = 126, p_8 = 126$$
$$g = 126, p_4 = 126, p_9 = 441, p_{10} = 336, p_7 = 336 \tag{A.20}$$

After exploiting all the possible symmetries, the total number of gain elements is 1820, for the specific example problem.

A.4 The Gain Differential Equations

Collection of coefficients required by the use of SSM yields a set of vector, matrix or tensor differential equations. Indicial notation development can also be exploited to bypass the usual rigorous process of generating the symbolic form of ordinary differential equations for the gains.

There are 15 distinct sets of the gain differential equations for the prototype problem,

i) Constant term:

$$-\dot{p}_0 = \frac{1}{2} L_{ij} p_i p_j + p_i C_i + C_J \tag{A.21}$$

where $L_{ij} \equiv -\frac{1}{2} r^{-1}_{mk} b_{im} b_{jk}$ (A.22)

ii) Linear term in x:

$$-\dot{p}_i = L_{jk}(p_j s_{ki} + p_k s_{ji}) + p_i a_{1_{ji}} + c_j s_{ji} + q_{1_i} \tag{A.23}$$

iii) Linear term in ν:

$$-\dot{p}_{1_i} = L_{jk}(p_j p_{5_{ki}} + p_k p_{5_{ji}}) + c_j p_{5_{ji}} \tag{A.24}$$

iv) Quadratic term in x (Unsymmetrical Riccati equation):

$$-\dot{s}_{jk} = q_{2_{jk}} + 2L_{im}(t_{ijk}p_m + s_{ij}s_{mk} + p_i t_{mjk}) + 2t_{ijk}c_i + 2s_{ij}a_{1_{ik}} + 2p_i a_{2_{ijk}} \quad \text{(A.25)}$$

Note that Eq. (A.25) is an unsymmetrical form of the Riccati equation,

v) Quadratic term in ν:

$$-\dot{p}_{2_{jk}} = 2L_{im}(p_{6_{jki}}p_m + p_{5_{ij}}p_{5_{mk}} + p_i p_{6_{jkm}}) + 2p_{6_{jki}}c_i \quad \text{(A.26)}$$

vi) Cross quadratic term in x and ν:

$$-\dot{p}_{5_{jk}} = L_{im}(2p_{8_{kij}}p_m + s_{ij}p_{5_{mk}} + 2p_i p_{8_{kmj}} + p_{5_{ik}}s_{mj}) + 2p_{8_{kij}}c_i + p_{5_{ik}}a_{1_{ij}} + 2p_i a_{2_{ijk}} \quad \text{(A.27)}$$

Now, all the third order terms can be collected as,

vii) $x_j x_k x_m \equiv$

$$-\frac{1}{3}\dot{t}_{jkm} = p_i a_{3_{ijkm}} + g_{ijkm}c_i + t_{ijk}a_{1_{im}} + L_{ie}\left(p_i g_{ejkm} + p_e g_{ijkm} + s_{ij}t_{ekm} + t_{ijk}s_{em}\right) + \\ + s_{ij}a_{2_{ikm}} + \frac{1}{3}q_{3_{jkm}} \quad \text{(A.28)}$$

viii) $x_j x_k \nu_m \equiv$

$$-\dot{p}_{8_{mjk}} = 2p_{8_{mik}}a_{1_{ij}} + 3p_{10_{mikj}}c_i + p_{5_{im}}a_{2_{ijk}} + L_{ie}(3p_i p_{10_{mejk}} + \\ + 3p_{10_{mikj}}p_e + 2p_{8_{mik}}s_{ej} + t_{ijk}p_{5_{em}} + p_{5_{im}}t_{ejk}) \quad \text{(A.29)}$$

ix) $\nu_m \nu_j x_k \equiv$

$$-\dot{p}_{6_{mjk}} = 2p_{6_{jmi}}a_{1_{ik}} + 2p_{9_{jmik}}c_i + L_{ie}(2p_i p_{9_{mjek}} + 2p_{9_{mjik}}p_e + p_{6_{mji}}s_{ek} + \\ + s_{ik}p_{6_{mje}} + 2p_{8_{jik}}p_{5_{em}} + 2p_{5_{ij}}p_{8_{mpk}}) \quad \text{(A.30)}$$

x) $\nu_j \nu_k \nu_m \equiv$

$$-\frac{1}{3}\dot{p}_{3_{jkm}} = p_{7_{jkmi}}c_i + L_{ie}(p_i p_{7_{jkme}} + p_{7_{jkmi}}p_e + p_{6_{jki}}p_{5_{em}} + p_{5_{ij}}p_{6_{kme}}) \quad \text{(A.31)}$$

Finally, the gain equation associated with the fourth-order terms can be given as,

xi) $x_j x_k x_m x_l \equiv$

$$-\frac{1}{4}\dot{g}_{jkml} = g_{ijkl}a_{1_{im}} + t_{ijk}a_{2_{iml}} + s_{ij}a_{3_{iklm}} + p_i a_{4_{imjkl}} + L_{ie}\left(t_{ijk}t_{elm} + s_{ij}g_{eklm} + s_{em}g_{ijkl}\right) + \frac{1}{4}q_{4_{jkml}} \quad \text{(A.32)}$$

xii) $x_j x_k x_m \nu_l \equiv$

$$-\dot{p}_{10_{ljkm}} = 3p_{10_{likm}}a_{1_{ij}} + 2p_{8_{lik}}a_{2_{imj}} + p_{5_{il}}a_{3_{ijkm}} + p_i a_{4_{imjkl}} + \\ + L_{ie}\left(3p_{10_{ljkj}}s_{em} + 2p_{8_{ljk}}t_{emj} + 2t_{ijk}p_{8_{lej}} + 3s_{ij}p_{10_{lekm}} + g_{ijkm}p_{5_{el}} + p_{5_{il}}g_{emjk}\right) \quad \text{(A.33)}$$

xiii) $x_j x_k \nu_m \nu_l \equiv$

$$-\dot{p}_{9_{jkml}} = 2p_{9_{mlik}}a_{1_{ij}} + p_{6_{mli}}a_{2_{ijk}} + \\ + L_{ie}\left(4p_{8_{lik}}p_{8_{mej}} + 2p_{9_{lmij}}s_{ek} + 3p_{5_{il}}p_{10_{mejk}} + p_{6_{mli}}t_{ejk} + t_{ijk}p_{6_{mle}} + 2s_{ij}p_{9_{mlek}} + 3p_{10_{likj}}p_{5_{em}}\right) \quad \text{(A.34)}$$

xiv) $x_j \nu_k \nu_l \nu_m \equiv$

$$-\dot{p}_{7_{klmj}} = L_{ie}\left(2p_{8_{eij}}p_{6_{klm}} + 2p_{6_{lki}}p_{8_{mej}} + p_{7_{mkli}}s_{ej} + s_{ij}p_{7_{mlke}} + 2p_{5_{ik}}p_{9_{lmej}} + 3p_{9_{lkij}}p_{5_{em}}\right) \quad \text{(A.35)}$$

xv) $\nu_j\nu_k\nu_l\nu_m \equiv$

$$-\frac{1}{4}\dot{p}_{4_{jklm}} = L_{ie}\left(p_{6_{jki}}p_{6_{lme}} + p_{5_{ij}}p_{7_{klme}} + p_{7_{jkli}}p_{5_{em}}\right) \quad \text{(A.36)}$$

All of these equations are given in their unsymmetrical form. To apply the available Runge-Kutta integration scheme with Matlab® 7.0, each gain set can be converted into a vector form by using the symmetry rules for each tensor defined above. After integrating the differential equations backwards in time, the gains can be restacked in the expansion of the value function by using multidimensional arrays and the rules for symmetry.

If the OCP is formulated with a soft constraint only, then there is no need for the terminal Lagrange multiplier, ν. A problem with hard constraint certainly requires ν. To evaluate ν, the series reversion process can be applied to Eq. (A.14). If the reversion is just considered up to second-order, ν is obtained as,

$$\nu = \alpha - \frac{1}{2}\beta + \frac{1}{6}\gamma \quad \text{(A.37)}$$

where,

$$\begin{aligned}
\alpha &= -(\bar{G}_1^{-1})(p_{1_i} + p_{5_{ji}}x_j + p_{8_{ijk}}x_jx_k + p_{10_{ijkl}}x_jx_kx_l) \\
\beta &= \bar{G}_1^{-1}(\bar{G}_2 \otimes \alpha \otimes \alpha) \\
\gamma &= \bar{G}_1^{-1}(\bar{G}_3 \otimes \alpha \otimes \alpha \otimes \alpha - 2\bar{G}_2 \otimes \beta \otimes \alpha - \bar{G}_2 \otimes \alpha \otimes \beta)
\end{aligned} \quad \text{(A.38)}$$

with

$$\begin{aligned}\bar{G}_1 &= p_{2_{ij}} + 2p_{6_{ijk}}x_k + 2p_{9_{ijkl}}x_k x_l \\ \bar{G}_2 &= 2p_{3_{ijk}} + 2p_{7_{ijkl}}x_l \\ \bar{G}_3 &= 6p_{4_{ijkl}}\end{aligned} \tag{A.39}$$

where, $\otimes$ represents the Kronecker product of two tensorial quantities.

The process can be extended in a similar way for implementing a higher order reversion. The appropriate boundary conditions for the gains can be obtained depending upon the presence of a terminal penalty function or a hard terminal constraint. The forward integration augmented with the series reversion process can be used to obtain the feedback control as well as the closed-loop state trajectories.

The following steps outline the process of collecting the gain differential equations:

1. Transform the plant equations and the performance index into the canonical forms given in Eqs. (A.4) and (A.5).
2. Provide the numerical values of the parameters, $a_{1_{ij}}, a_{2_{ijk}}, a_{3_{ijkl}}, a_{4_{ijklm}}, C_i$ and b_{ij} in the appropriate tensorial form.
3. Stack all the tensors in multidimensional arrays appropriately by using the symmetry properties.

The main advantage of this procedure is to improve upon the tedious task of symbolic computation performed by collecting the gain coefficients individually. Another advantage is that the gains can be programmed for a general class of problems.

APPENDIX B

CHEBYSHEV POLYNOMIALS

In this Appendix, the Chebyshev polynomial expansions utilized in GAT are provided.

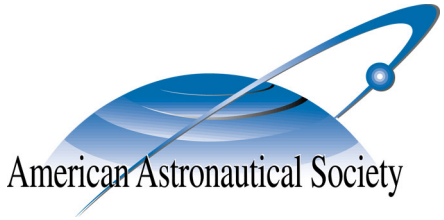
B.1 Orthogonal Bases Functions (Chebyshev Polynomials of first kind) upto 8th Order

$$T_0(x) = 1$$
$$T_1(x) = x$$
$$T_2(x) = 2x^2 - 1$$
$$T_3(x) = 4x^3 - 3x$$
$$T_4(x) = 8x^4 - 8x^2 + 1$$
$$T_5(x) = 16x^5 - 20x^3 + 5x$$
$$T_6(x) = 32x^6 - 48x^4 + 18x^2 - 1$$
$$T_7(x) = 64x^7 - 112x^5 + 56x^3 - 7x$$
$$T_8(x) = 128x^8 - 256x^6 + 160x^4 - 32x^2 + 1$$

The weight function for the Chebyshev polynomial of first kind is given as, $w(x) = \frac{1}{\sqrt{1-x^2}}$.

APPENDIX C

COPYRIGHT PERMISSIONS



6352 Rolling Mill Place, Suite 102
Springfield, VA 22152
Phone: 703-866-0020; Fax: 703-866-3526
aas@astronautical.org
www.astronautical.org

February 7, 2008

Copyright Permission

The American Astronautical Society grants copyright permission to Rajnish Sharma for the paper AAS 05-182: *Optimal Finite-Time Feedback Controllers for Nonlinear Systems with Terminal Constraint*, published in the Proceedings of the 2005 AAS/AIAA Spaceflight Mechanics Winter Meeting, Vol. 120, Part 2, pp. 1317-1330.

James R. Kirkpatrick
Executive Director

6352 Rolling Mill Place, Suite 102
Springfield, VA 22152
Phone: 703-866-0020; Fax: 703-866-3526
aas@astronautical.org
www.astronautical.org

February 7, 2008

Copyright Permission

The American Astronautical Society grants copyright permission to Rajnish Sharma for the paper AAS 05-337: *Higher-Order Neighboring-Optimal Guidance for Continuous-Thrust Orbit Transfer*, published in the Proceedings of the 2005 AAS/AIAA Astrodynamics Specialist Conference, Vol. 123, pp. 1321-1336.

James R. Kirkpatrick
Executive Director

6352 Rolling Mill Place, Suite 102
Springfield, VA 22152
Phone: 703-866-0020; Fax: 703-866-3526
aas@astronautical.org
www.astronautical.org

February 7, 2008

Copyright Permission

The American Astronautical Society grants copyright permission to Rajnish Sharma for the paper AAS 07-384: *Optimal Strategies in Cooperative Games with Terminal Payoff*, published in the Proceedings of the 2007 AAS/AIAA Astrodynamics Specialist Conference.

James R. Kirkpatrick
Executive Director

6352 Rolling Mill Place, Suite 102
Springfield, VA 22152
Phone: 703-866-0020; Fax: 703-866-3526
aas@astronautical.org
www.astronautical.org

February 7, 2008

Copyright Permission

The American Astronautical Society grants copyright permission to Rajnish Sharma for the paper AAS 07-411: *A Series Solution Method for the Solution of the Hamilton Jacobi Isaacs Equation and Its Applications to Aerospace Systems*, published in the Proceedings of the 2007 AAS/AIAA Astrodynamics Specialist Conference.

James R. Kirkpatrick
Executive Director

6352 Rolling Mill Place, Suite 102
Springfield, VA 22152
Phone: 703-866-0020; Fax: 703-866-3526
aas@astronautical.org
www.astronautical.org

February 7, 2008

Copyright Permission

The American Astronautical Society grants copyright permission to Rajnish Sharma for the paper AAS 07-412: *Optimal Nonlinear Feedback Controller Design Using a Waypoint Scheme*, published in the Proceedings of the 2007 AAS/AIAA Astrodynamics Specialist Conference.

James R. Kirkpatrick
Executive Director

VITA

Rajnish Sharma received his Bachelor of Technology in mechanical engineering from Indian Institute of Technology, Kanpur in 1999. He entered the Aerospace Department at Indian Institute of Technology, Kanpur in July 2001 and received his Master of Technology degree in May 2003. He received his doctoral degree in Aerospace Engineering from Texas A&M University, College Station in 2008. His research interests include optimal control theory, differential games and nonlinear dynamical systems. He plans to publish a book on these topics, focusing on nonlinear feedback methods by using the series solution approach.

Rajnish Sharma may be reached at the Department of Aerospace Engineering and Mechanics, the University of Alabama, Tuscaloosa, AL-35406. His email is rsharma@eng.ua.edu.